A genetic approach to identify *Hox* regulatory microRNAs during *Drosophila* development

Wan Liu

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Sussex

September 2015

UNIVERSITY OF SUSSEX

Wan Liu, DPhil Biology

A genetic approach to identify *Hox* regulatory microRNAs (miRNAs) during *Drosophila* development

Summary

The *Hox* genes encode a family of transcriptional regulators that activate distinct developmental programs along the anterior-posterior (AP) axis of animals. Recent observations in *Drosophila* demonstrate that at least two miRNAs can repress *Hox* gene expression during development suggesting that miRNA-based regulation might be a general mechanism of *Hox* gene regulation.

Here explore this possibility by applying a comprehensive genetic approach to identify miRNAs able to repress *Hox* gene expression during development. Given that the reduction of *Drosophila Hox* gene *Ultrabithorax* (*Ubx*) expression leads to easily tractable homeotic transformations in haltere, I use *Ubx* to test the repressive effects of dozens of miRNAs in an overexpression screen.

Scoring over 10,000 halteres showed that out of 106 miRNAs tested, ~28% produced *Ubx* mutant phenotypes suggesting that miRNA-dependent *Hox* regulation might be a pervasive mechanism controlling *Hox* gene function during development. I classify phenotypes into four major categories: *Ubx* mutant effects (Class I and II) and others (Class III and IV).

Through the combination of RNA-Seq data and TaqMan RT-PCR approaches, I confirm that there is no correlation between the phenotypic strength and miRNA expression level indicating that haltere phenotypes emerge from miRNA qualitative roles. Furthermore, using protein expression analysis and *Ubx* 3' UTR fluorescent reporters, I confirmed that at least nine miRNAs affect Ubx protein expression and that six of these directly target *Ubx* 3' UTR *in vivo*.

Lastly, I explore the nature and effects of miRNA regulation of *Ubx* at the cellular level in the *Drosophila* embryonic CNS and find that *miR-252* is sufficient and necessary to repress *Ubx* expression in specific neural lineages.

Our work thus contributes to the understanding of miRNA-mediated *Hox* gene regulation and, more generally, to the study of miRNA-target interactions within the physiological context of metazoan development.

Acknowledgements

First, I would like to thank my supervisor, Claudio, for his persistent encouragement and guidance, for all the discussions of my project and this thesis, for getting me to think for myself to become a more independent researcher. I want to thank my co-supervisor Majid, who gave me lots of support and help during my study here.

I also want to thank all past and present members of the Alonso group- Casandra, Rich, João, Pedro, Ipek, Rual, Ines, Sofia and Aalia, who provided a fun, relaxed environment to complete this work. Special thanks must also go out to Pedro, João, Rual and Ines for all the discussions and help, and especially the support in the final days of completing this thesis.

Thanks to everyone in the fly groups, especially Inyaki, Unum, Ali, Emile, Ying, Will, John and Chris, and the badminton group, Julie, Ilona, Tanya and Sarah. I would like to thank Chris for teaching me the basic setup of TaqMan RT-PCR and Mansi for showing me how to use the machine. Thanks to Dan, Roger, Julian and Sofia for their technical support.

I would like to thank all friends I met here, especially Sunny, Lina. We share almost everything together and support each other whenever we need help. Also, I would like to thank Andy, Chris, Eva, Antonio, Tali, Emyre, Sheng, Xiaoxiao, Kwok, Ana, Cuihua, Yifan, Xuejuan, Yuyu, as well as friends at home: Yan, Jinsong, Ruoyang, Tingting, Jie, Chaowei, Xiaoge, Beibei and all my friends whose name are not here. Special thanks must go out to Yongqing for the support of bioinformatic analysis. A big thank you to Hongliang for all the discussion, support and help.

I would like to thank the China Scholarship Council and University of Sussex for providing me the funding to pursure my PhD project. I also want to thank Great Britain-China Education Trust and Welcome Trust.

A massive thank you to all my family, especially my mum, my dad, my grandparents and my sister. This thesis is dedicated to you.

To all of you, a very, very big thank you.

To my family

Table of Contents

List of Figures

List of Tables

Abbreviations

ac/sc – achaete scute

APA – alternative polyadenylation

abd-A – abdominal-A

Abd-B – Abdominal-B

AGO – Argonaute

Antp – Antennapedia

ANT-C – Antennapedia complex

AP – anterior-posterior

a.u. – arbitrary fluorescence intensity units

BH – bigger haltere

BMP – bone morphogenetic protein

BX-C – Bithorax complex

cas – castor

CDS – coding sequence

C.elegans – Caenorhabditis elegans

CNS – central nervous system

Dfd – Deformed

DV – dorsal-ventral

EB – ectopic bristles

EGFR – epidermal growth factor receptor

FPKM – fragments per kilobase of transcript per million fragments mapped

GMC – ganglion mother cells

GOF – gain-of-function

grh – grainy head

hb – hunchback

ind – intermediate neuroblasts defective

kb – kilobase

Kr – Krüppel

LOF – loss-of-function

lab – labial

miRNA – microRNA

miR-310C – miR-310-313 cluster

MREs – miRNA recognition elements

msh – *muscle segment homeobox*

NGS – next generation sequencing

Or-R – Oregon Red

PAR-CLIP – photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation

PcG – *Polycomb* group

pb – *proboscipedia*

PS – parasegment

pre-miRNA – precursor microRNAs

pri-miRNA – primary microRNA

RBP – RNA binding protein

RISC – RNA-induced silencing complex

RNAP II – RNA polymerase II

RNA-seq – RNA sequencing

Scr – *Sex combs reduced*

SEM – scanning electron microscope

SH – small haltere

TRBP – TAR RNA-binding protein

TrxG – *Trithorax* group

Ubx – *Ultrabithorax*

UTR – untranslated region

VNC – ventral nerve cord

vnd – *ventral nervous system defective*

WPP – white pre-pupae

3’ UTRs – 3’ untranslated regions

5’ UTRs – 5’ untranslated regions

CHAPTER 1 Introduction

1.1 Preface

How is genomic information converted into the specific cellular programs that underlie morphology? This is in part achieved through the differential activation of genes at particular temporal and spatial coordinates. Different layers of gene regulation contribute to the specification of cell identities, including transcriptional and post-transcriptional modulation (Orphanides & Reinberg 2002; Alonso & Wilkins 2005; Alonso 2008). The work presented here focuses on this later regulation level.

Increasing evidence indicates that post-transcriptional regulation is more important in the process of gene expression than previously thought (Alonso & Wilkins 2005). For example, the *Drosophila* let-7-Complex controls the transition of larval-to-pupal stage through regulating the transcription factor *chinmo* in specific neuronal lineages in the *Drosophila* brain (Wu et al. 2012). microRNAs (miRNAs) are typical post-transcriptional regulators that determine the functional outcome of an mRNA. By binding to the target mRNA in a complementary manner (Figure 1.1), miRNAs manipulate mRNA expression levels by either repressing mRNA translation or mediating mRNA degradation (Bartel 2009; Filipowicz et al. 2008). It has been shown that miRNAs regulate more than 30% human protein coding genes (Bartel 2004; Eulalio et al. 2008; Filipowicz et al. 2008; Kloosterman & Plasterk 2006; Nilsen 2007). However, the mechanisms that allow miRNAs to identify their targets in the context of development are still not fully understood. This makes it difficult to determine the full repertoire of active miRNAs controlling gene expression profiles within the dynamic context of development.

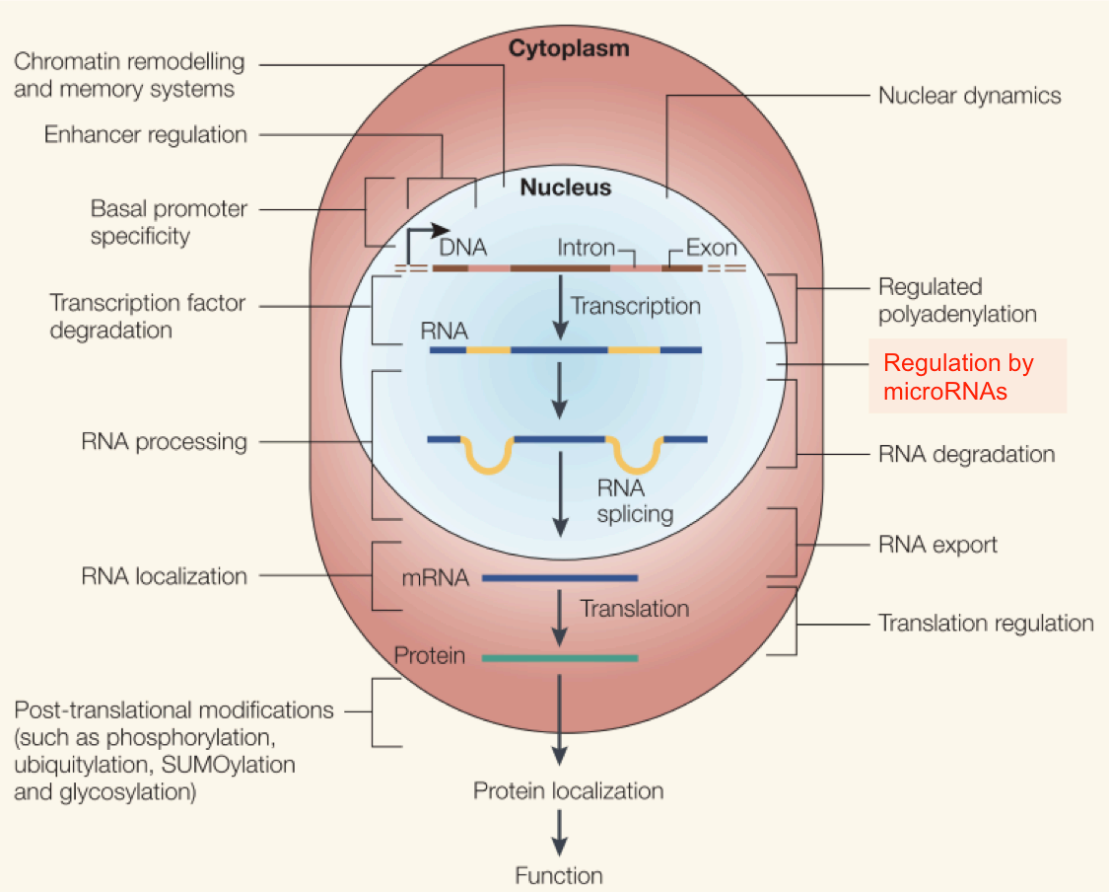


Figure 1.1 Multiple layers of gene regulation.

Diagram shows the different layers of gene regulation. Post-transcriptional regulation is a key step in the control of gene expression at RNA level, including several processes such as splicing, RNA modification, stability, polyadenylation, localization and translation. Current studies have shown that miRNA-mediated post-transcriptional regulation is more important than previously acknowledged (Krol et al. 2010). Figure adapted from (Alonso & Wilkins 2005).

The *Hox* genes encode a family of transcriptional regulators that activate distinct developmental programs along the AP axis of animals (Pearson et al. 2005). At early development, all cells within each *Drosophila* "segment" are allocated a unique '*Hox code*' (Wellik 2007). An important question in the field relates to the specific regulatory steps that pattern *Hox* activities during *Drosophila* development e.g. how do miRNAs regulate *Hox* expression patterns post-transcriptionally. In *Drosophila,* it has been reported that two miRNAs, *dme-miR-iab-4* and *dme-miR-iab-8* fine-tune the expression of the developmental gene *Ubx* (Bender 2008; Ronshaugen et al. 2005; Stark et al. 2008; Thomsen et al. 2010; Tyler et al. 2008). This study will focus on uncovering the full miRNA repertoire affecting the expression of *Ubx*, in an

attempt to understand the miRNA-dependent *Hox* gene regulation during development.

In the present work, firstly, I perform a large-scale genetic screen to uncover potential miRNAs that may regulate *Ubx* expression in *Drosophila*. This was followed by an analysis of native miRNA expression levels for all screened miRNAs to exclude potential biases in the screen. Secondly, I validated miRNAs that directly target *Ubx* 3' UTRs *in vivo*. Exploiting different miRNA target-prediction softwares, I propose target sites for all six miRNAs that directly target *Ubx* 3' UTRs. Finally, I explored the regulatory potential of a subset of miRNAs identified in the genetic screen in regard with *Hox* gene regulation within the developing embryonic central nervous system (CNS).

1.2 The *Hox* genes

Mutations in *Hox* genes led to homeotic transformations: the morphological transformation of one body part into the likeness of another (Bateson 1894). In 1923, Calvin Bridges and Thomas Hunt Morgan isolated and identified the first *bithorax* mutation (Bridges & Morgan 1923). Following these initial studies, Lewis reported that the *Bithorax* complex (BX-C) regulates cell determination in *Drosophila* (Lewis 1951). In 1978, Edward B. Lewis published the chromosomal maps of the BX-C locus, and described a series of BX-C mutants that cause a haltere-to-wing transformation in line with a homeotic change in appendage specification (Figure 1.2) (Lewis 1978).

Hox genes are members of a large gene family, characterised by the presence of the homeobox, a 180 nucleotide-long DNA sequence that encodes for a 60 amino acids DNA-binding domain, known as the Homeodomain (Banerjee-Basu & Baxevanis 2001). The proteins encoded by *Hox* genes, called Hox proteins, bind to DNA and function as transcription factors to regulate the transcriptional activity of downstream genes (Pearson et al. 2005). *Hox* genes are determinant factors to the basic structure of the organism (Hirth et al. 1998).

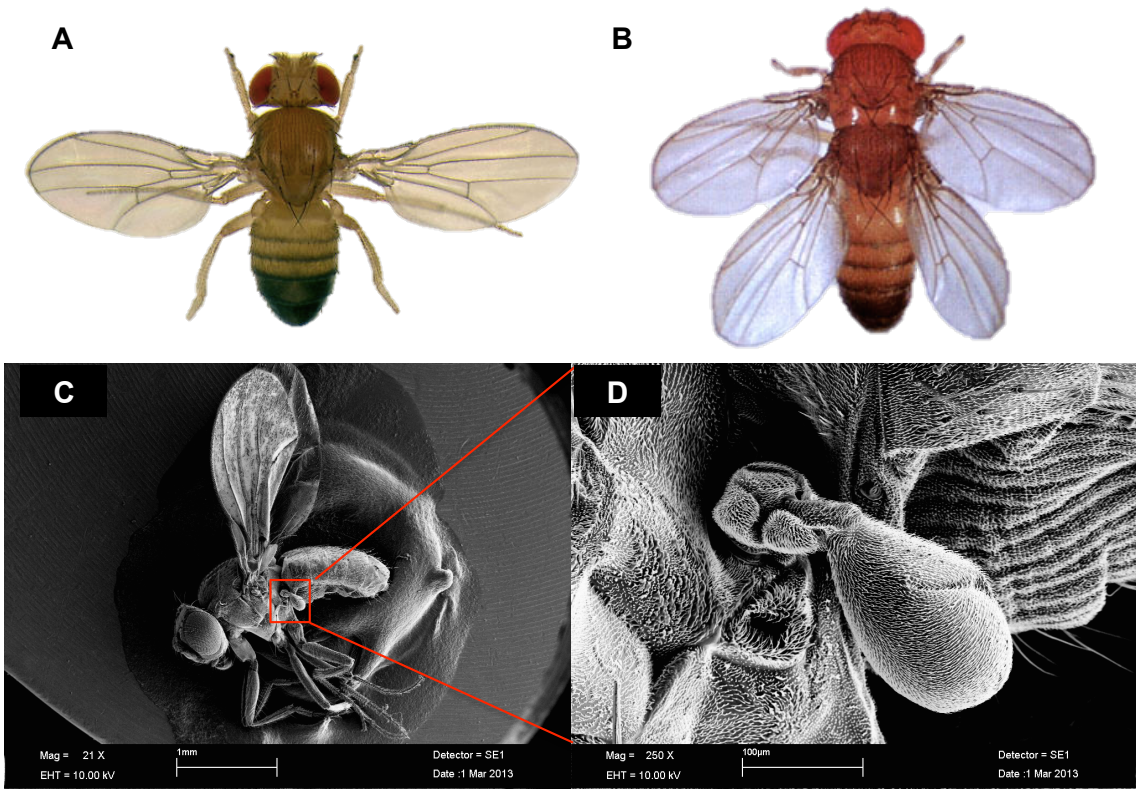


Figure 1.2 Haltere-to-wing transformation.

Dorsal view of (A) wild-type fly and (B) *Ubx* mutant fly (Photoes taken by E.B.Lewis). *Ubx* mutation shows haltere-to-wing transformation. (C) SEM image of a wild-type adult fly. (D) Enhanced view of the wild-type adult haltere.

Hox gene expression obeys a certain spatial order defined by the phenomenon of collinearity; where the sequential positioning of *Hox* genes within their genomic cluster corresponds to their sequential expression patterns along the AP axis (Mallo & Alonso 2013). *Hox* genes are composed of two separate gene clusters in *Drosophila* (Figure 1.3A): the *Antennapedia* complex (ANT-C) (Lewis et al. 1980a; Lewis et al. 1980b; Kaufman et al. 1980) and the *Bithorax* complex (BX-C) (Lewis 1978). The ANT-C contains *Hox* genes that regulate the identity of anterior structures during *Drosophila* development including the genes of *labial* (*lab*), *proboscipedia* (*pb*), *Deformed* (*Dfd*), *Sex combs reduced* (*Scr*) and *Antennapedia* (*Antp*), while BX-C controls the development of structure and function between the mid-metathorax and the more posterior abdominal segments (Gaunt 2015) containing the genes of *Ubx*, *abdominal-A* (*abd-A*) and *Abdominal-B* (*Abd-B*). The *Hox* genes are evolutionary conserved between invertebrates and vertebrates (W McGinnis et al. 1984a; W McGinnis et al.

1984b). The invertebrates, such as *Drosophila*, have homologous counterparts *Hox* genes in vertebrates (Figure 1.3B). However, while the *Hox* genes in *Drosophila* are separated into two complexes, within a single *Hox* cluster, the vertebrate *Hox* genes have undergone two rounds of duplication that resulted in four clusters, a process that is thought to be related to the complexity of vertebrate body plan (Meyer 1998).

The conservation in genomic arrangement, colinear expression patterns and developmental functions of *Hox* genes across evolution point to the importance of *Hox* genes in shaping the body plan within and across species (Carroll 1995; Hughes & Kaufman 2002). Within one given species, different *Hox* genes provide positional identity to different body regions. For example, ANT-C regulates anterior of *Drosophila* body, while BX-C modulates the posterior region (Gaunt 2015). The *Hox* gene *Ubx* shapes appendage morphology in *Drosophila* and water striders (Stern et al. 1998; Khila et al. 2009). Among different species, the homologous *Hox* genes define the body region in a homologous position in different ways. For instance, the expression of *Hox* gene *Ubx* differentiates the hindwing from the forewing in different insects (Beeman et al. 1993; Warren et al. 1994; White & Wilcox 1984; Beachy et al. 1985).

Abnormal *Hox* genes expression causes developmental problems and may also lead to disease (Quinonez & Innis 2014; Cillo et al. 1999; Boncinelli 1997; Procino & Cillo 2013; Bene & Wittbrodt 2005; Sun et al. 2013; Raman et al. 2000; Muragaki et al. 1996; Pietro et al. 2012). For example, a missense mutation in the homeobox of *HoxA13* and a dinucleotide deletion in the promoter cause the Guttmacher syndrome, which is a dominantly inherited combination of distal limb and genital tract abnormalitites (Innis et al. 2002). Many mechanisms, such as chromatin remodeling, RNA processing, miRNA regulation and translation regulation, have been suggested to play roles in controling *Hox* gene expression (Alonso & Wilkins 2005; Mallo & Alonso 2013). In spite of crucial roles in body patterning, only a limited number of *Hox* controlled morphogenetic and differentiation patterns have been studied in detail (Alonso 2002; Hersh et al. 2007). A closer study of the mechanism that

controls *Hox* expression and function in *Drosophila* may elucidate mechanisms that cause *Hox* related developmental problems and diseases.

In this study, I used *Hox* gene *Ubx* to understand the mechanisms that allow miRNAs to identify their targets in the context of development. Particularly, I have focused on uncovering the miRNA repertoire affecting the expression of *Ubx*.

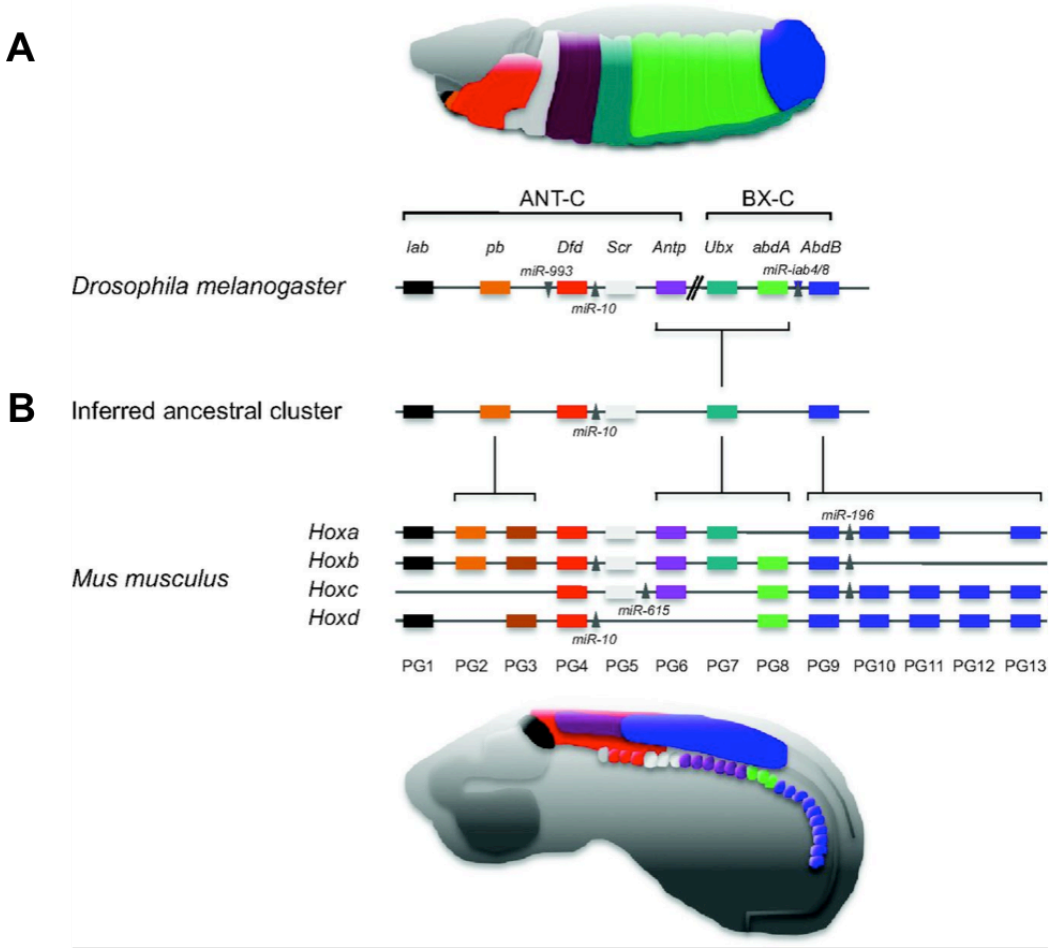


Figure 1.3 Hox genes are evolutionary conserved between invertebrates and vertebrates.

Taken from Mallo & Alonso (Mallo & Alonso 2013). (A) In *Drosophila*, *Hox* genes are composed of two separate gene clusters: ANT-C and BX-C. The ANT-C, which includes *lab*, *pb*, *Dfd*, *Scr* and *Antp*, contains *Hox* genes that regulate the identity of anterior structures during *Drosophila* development, while BX-C, consisting of *Ubx*, *abd-A* and *Abd-B*, controls the development of structures and function between the mid-metathorax and the more posterior abdominal segments. (B) The *Hox* genes are evolutionary conserved between invertebrates and vertebrates. As such, the invertebrate *Hox* genes have homologous counterparts in vertebrates, which also display spatial collinearity between genomic locus and gene expression during development. *Drosophila* miRNAs, *miR-993*, *miR-10* and *miR-iab-4/8* denoted by triangles are miRNAs encoded within the *Hox* cluster.

1.3 The *Drosophila Hox* gene *Ubx*

The *Drosophila Hox gene Ubx* occupies a genomic DNA stretch of around 78 kilobase (kb) pairs within the BX-C. This locus controls developmental processes that establish the adult morphology of regions between the posterior portion of second thoracic and the anterior part of the eighth abdominal segments. The mRNA and protein expression levels of *Ubx* during embryogenesis are very well documented (White & Wilcox 1985; Akam & Martinez-Arias 1985; Irvine et al. 1991; Beachy et al. 1985). *Ubx* transcription is first detectable in the developing blastoderm (Akam & Martinez-Arias 1985) and Ubx protein expression is first observed in the epidermis of parasegment (PS) 6 at early stage 9 (Irvine et al. 1991). During embryogenesis, Ubx is restrictedly expressed in the CNS, epidermis, somatic mesoderm and visceral mesoderm (White & Wilcox 1985). In the *Drosophila* CNS, *Ubx* expression occurs across PS5 to PS13 after germ band retraction. Ubx has the highest expression level in PS6, with a lower expression level in PS5, and gradually declines in expression from PS6 to PS13. Moreover, Ubx expression levels can vary immensely between individual cells within a single PS (White & Wilcox 1984; Irvine et al. 1991).

Beachy and colleagues have observed that Ubx expression in imaginal discs follows the aforementioned anterior to posterior embryonic expression patterns (Beachy et al. 1985). Ubx is strongly expressed in haltere imaginal disc and third leg imaginal disc in 3rd thoracic segment (T3), while the wing disc in T2, Ubx is only expressed in the peripodial membrane (White & Wilcox 1984). Ubx expression is higher in the posterior part of both haltere and third leg imaginal discs (White & Wilcox 1985). In addition, *Ubx* has also been shown to control the morphology of halteres and legs during *Drosophila* development (Roch & Akam 2000; Navas et al. 2011).

How are these precise *Ubx* patterns established in *Drosophila* development? Four parasegment-specific cis-regulatory domains of BX-C, *abx*, *bx*, *bxd* and *pbx*, in *abx/bx* and *bxd/pbx* regions are involved in the control of *Ubx* transcription (Little et al. 1990; Beachy et al. 1985; Maeda & Karch 2006).

Abx/bx and *bxd/pbx* control *Ubx* expression in parasegments 5 (PS5) and PS6, respectively. The initial boundaries of *Ubx* expression are set by the maternal, gap genes and pair-rule genes (White & Lehmann 1986; Ingham & Martinez-Arias 1986; Reinitz & Levine 1990; Kornberg & Tabata 1993). Nevertheless, the gap and pair-rule genes are only transiently expressed; therefore, another system that maintains the *Ubx* expression during development is needed. The products of the epigenetic regulators, *Polycomb* group (PcG) and *Trithorax* group (TrxG) genes are required for this maintenance system. The *Drosophila* PcG and TrxG act antagonistically to maintain *Ubx* silenced and activated by modifying the chromatin structure of the *Ubx* cis-regulatory regions (Pirrotta 1997; Paro 1990; Simon 1995; Kennison 1993; Steffen & Ringrose 2014). The PcG proteins maintain the inactive state of *Ubx* cis-regulatory region, whereas the TrxG proteins keep the active state of *Ubx* cis-regulatory region, to ensure that the initial transcription of *Ubx* expression can be precisely regulated throughout *Drosophila* development.

Ubx expression is firstly established by the segmentation cascade genes by acting on specific regulatory regions, and secondly its expression is maintained throughout development by epigenetic factors. Thus, *Ubx* expression is able to regulate the formation of different organs throughout development, such as appendages and CNS.

1.4 *Hox* regulation during haltere development

Appendage development relies on coordinated gene regulations, such as the regulations of transcription factors and signaling systems, which act as crucial components in the genetic programs that control appendage development (Lecuit et al. 1996; Blair et al. 1994; Morata 2001; Pearson et al. 2005). For example, *Hox* genes control the specification of particular bristle types and cell shapes within the *Drosophila* appendages (Rozowski & Akam 2002; Roch & Akam 2000; Kaufman et al. 1980) and the development of the underlying musculature and neuronal connections that innervate the appendage (Fernandes et al. 1994; Burt & Palka 1982).

In *Drosophila*, the haltere is a small dorsal appendage in the third thoracic segment, which is involved in the control of flight (Morata 2001). Interestingly, the development of the *Drosophila* haltere is completely dependent on the regulation of the *Hox* gene *Ubx*. Homeotic transformation from haltere to wing commonly occur in *pbx* mutant flies (Lewis 1978), conversely, *Ubx* gain-of-function mutations can lead to wing-to-haltere transformations (Pavlopoulos & Akam 2011). The aforementioned *Ubx* gain-of-function and loss-of-function observations are consistent with the argument that *Ubx* governs haltere development. In the following paragraphs, I will focus on the molecular mechanism by which *Hox* gene *Ubx* controls *Drosophila* haltere development.

Ubx shapes haltere morphology through multiple levels of regulatory hierarchies (Weatherbee et al. 1998). *Ubx* acts independently on at least five genes that are themselves regulators in the wing patterning hierarchy. For example, *Ubx* independently represses the expression of the *Serrate* (*Ser*) and *wingless* (*wg*) signal in the posterior region of the developing haltere imaginal disc as well as the *wg* downstream genes, *AS-C* and *vestigial* (*vg*) (more specifically the quadrant enhancer), along the dorsoventral (DV) boundary (Weatherbee et al. 1998). In addition, *Ubx* also regulates *Vg* downstream gene, *Drosophila Serum Response Factor* (*DSRF* or *blistered*) in the proximodistal axis. Along the anteroposterior boundary, *Ubx* independently controls *decapentaplegic* (*dpp*) and its downstream gene *spalt-related* (*salr*) (Navas et al. 2006; Weatherbee et al. 1998). The proximal-distal anatomy of the *Drosophila* haltere consists of three parts: the scabellum, the pedicellus and the capitellum (Roch & Akam 2000). The pouch at the centre of the haltere imaginal disc develops into the capitellum, while the surrounding fold will form the pedicellus (Figure 1.4A). During haltere imaginal disc development, *Ubx* controls the development of each portion of the tissue through distinct genetic programs using the genes in the *Ubx* regulatory hierarchies (Held 2002). *Ubx* represses the expression of *blistered* in the haltere to form the balloon-like structure, the capitellum. By independently repressing *spalt*, *salr*, *vg* and *wg*, *Ubx* shapes the size of the haltere from the ground state (wing). To repress the formation of bristles along the wing margin, a main feature of *Drosophila* wings, *Ubx* represses the expression of *AS-C* and *wg* in the capitellum, while activating *AS-C* to promote

the differentiation of stretch-sensitive sensilla at the base of the pedicel (Weatherbee et al. 1998).

Ubx regulation of haltere morphology relies on the dosage of *Ubx*. It has been demonstrated that increasing *Ubx* expression in the haltere imaginal disc leads to reduced haltere sizes (Smolik-Utlaut 1990; Crickmore et al. 2009). In contrast, different *Ubx* mutant alleles (Figure 1.4B) are characterized by different *Ubx* expression levels in the haltere imaginal discs, which are correlated with the varying severities of haltere-to-wing homeotic transformations in the adult haltere appendage (Bender et al. 1983). Altogether, the evidence above suggests that developing haltere imaginal disc are sensitive to *Ubx* expression level changes, and that these changes can lead to significant alterations in haltere morphology.

Ubx plays important roles in the development of animal appendages at the cellular level. *Ubx* expression is required during the latest stages of larval development to shape cell morphology in order to form the haltere (Roch & Akam 2000). At the end of pupal development, haltere cells are 8-fold smaller in terms of apical surface area compared to wing cells, and display multiple trichomes per cell, while cells in the wing only have a single bristle. This is consistent with previous argument that instead of controlling segment identity, *Hox* genes control cellular behavior that will result in a certain segment morphology (Castelli-Gair 1998). However, the mechanism underlying this kind of cellular-level haltere morphology regulation works is not fully understood. In terms of studying the mechanism of cellular-level regulation, the cell lineages in the CNS are much well studied compared to the haltere, thus, CNS is a much bettern system (Bossing et al. 1996; Schmidt et al. 1997). I will describe in detail in the next section.

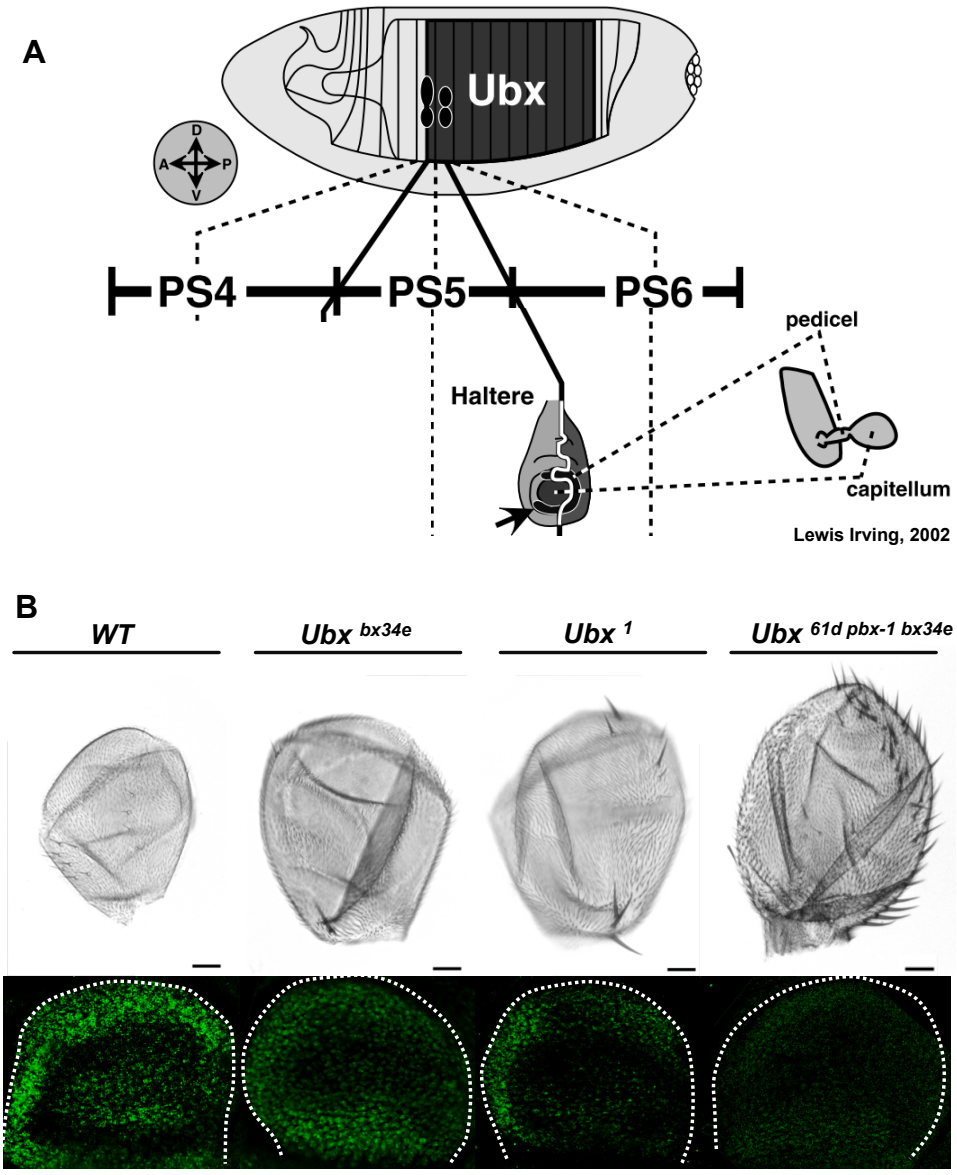


Figure 1.4 *Ubx* regulation of haltere morphology relies on the dosage of Ubx protein.

(A) Fate map of the left half of a stage 11 embryo shows where Ubx is expressed in the ectodermal regions of parasegment (PS) 5-13. A-P, anterior-posterior; D-V, dorsal-ventral. The cell lineages developed to change for the ectodermal primordial (circles) that give rise to the haltere imaginal disc are located between the posterior of PS5 and the anterior of PS6 in the *Drosophila* embryo. The pouch at the centre of the haltere imaginal disc develops into the capitellum, while the surrounding fold will form the pedicel. The degrees of shading in the haltere imaginal discs denotes the different amounts of Ubx expression levels. The Ubx expression in the haltere imaginal discs is generally stronger in the posterior compartment than the anterior compartment, but the intensity is also high in the part of anterior region (arrow) near the center and in the surrounding fold. Taken from Lewis Irving Held (Held 2002) (B) An allelic series of *Ubx* mutations showing an increase in phenotypic severity from left to right. Changes in phenotype correlate with gradual loss of Ubx expression within the pouch region of the haltere imaginal discs. The *Ubx* allelic series includes the following genotypes: *w;;bx^{34e}/bx^{34e}*, *w;;Ubx^{1}/TM6b* and *w;;Ubx^{61d} pbx^{1}/bx^{34e}* in increasing order of severity, taken from Richard Kaschula (Richard Kaschula PhD thesis; The regulation of *Hox* genes by *microRNAs* during *Drosophila* development, 2013; Fig, 1.7D).

1.5 *Hox* regulation during CNS development

CNS is a component of the nervous system, which processes sensory information from the whole body and provides a corresponding response in bilaterian animals such as *Drosophila*, mouse and human (Holland et al. 2013; Kourakis et al. 1997). A large number of genes and cellular processes are tightly regulated in time and space to coordinate CNS development in *Drosophila* (Technau et al. 2006; Skeath & Thor 2003). The insect CNS contains two main anatomical parts: brain and ventral nerve cord (VNC) (Demerec 1950; Hartenstein 1993). As *Ubx* is expressed in the VNC, this work focuses on the *Drosophila* VNC development.

The VNC is developed from neuroectodermal cells in the ventral-lateral regions of the embryo in *Drosophila*, and patterning of the neuroectoderm occurs at early stages of embryogenesis along the AP and DV axes of the embryo (Figure 1.5A), producing neural equivalence group by a Cartesian coordinate system (Figure 1.5B). The gap and pair-rule genes define the segment-polarity genes expression within each segment in the VNC along the AP axis, and the segment-polarity genes pattern and define the fate of neuronectodermal cells in each row (Akam 1987; Bhat 1999). The combination of the functions of the nuclear factor NF-κB, bone morphogenetic protein (BMP) and the epidermal growth factor receptor (EGFR) along the DV axis regulate the expression of ventral nervous *system defective* (*vnd*), *intermediate neuroblasts defective* (*ind*) and *muscle segment homeobox* (*msh*), dividing the neuroectoderm into three longitudinal stripes, ventral, intermediate and dorsal columns (Skeath 1998; von Ohlen & Doe 2000).

After gastrulation, the neuroectoderm cells undergo either neurogenesis or epidermogenesis. The neurogenesis of the VNC consists of two main steps. First, within each hemisegment, the ventral neurogenic ectoderm delaminates into the embryo, and becomes a stereotyped subepidermal array of neuronal precursor cells or neuroblasts (NBs) (Doe et al. 1988; Jiménez & Campos-Ortega 1990). The combination of *achaete scute* (*ac/sc*) Complex and Notch/Delta signaling pathway (Heitzler et al. 1996; Heitzler & Simpson 1991),

determines the selection of one cell within the proneural cluster (Figure 1.5C). All cells in the proneural cluster have the same potential to develop as a NB, and the expression of *ac/sc* can promote the NB fate (Jiménez & Campos-Ortega 1990). *ac/sc* activates *Delta* expression, which activates the Notch in the neighboring cell, thereby repressing *ac/sc* expression there (Skeath & Carroll 1992). Thus, one proneural cell initially with a relative higher expression of ac/ac or Delta and lower expression of Notch, has the ability to activate the Notch signaling in the neighboring cells and repress *ac/sc* expression. Then this cell can develop as NB, and migrate to the interior of the embryo. In each hemisegment, about 30 NBs are produced after five sequential waves of NB segregation (Figure 1.5D).

Second, a three-dimensional CNS structure progressively results from the transformation of the two-dimensional layer of NBs. Each NB goes through an invariant stem cell lineage, which forms ganglion mother cells (GMC) into the embryo (Thomas, Bastiani et al. 1984; Thomas, Crews et al. 1988). The specification of GMCs is controlled by sequential expression of five genes, *hunchback* (*hb*), *Krüppel* (*Kr*), *Pdm*, *castor* (*cas*) and *grainy head* (*grh*) (Figure 1.5E). These five genes are transcription factors that expressed in a temporal cascade in each NB, and then transmitted to the GMC and postmitotic cells. Under the control of the antagonistic interactions between the Notch pathway and Numb, the asymmetric division of GMC specifies the different identities of glial and neuronal cells (Figure 1.5F). About 30 glia and 350 neurons are produced in each hemisegment of the VNC at the end of embryogenesis (Schmidt et al. 1997). The diverse arrangement of neurons and glia in the embryonic CNS forms the complexity of nervous system.

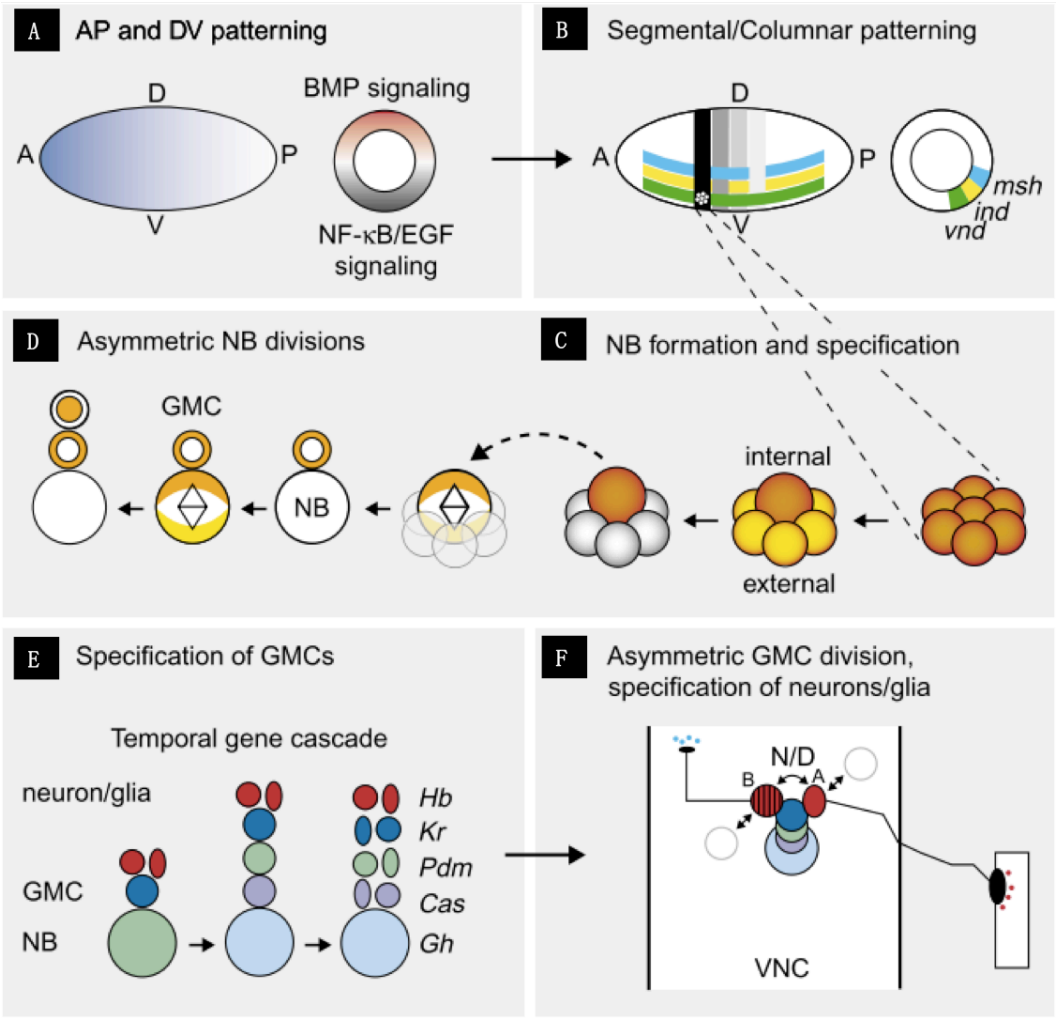


Figure 1.5 Development of *Drosophila* VNC during embryogenesis.

Taken from Skeath and Thor (Skeath & Thor 2003). In *Drosophila*, the VNC is developed from neuroectodermal cells in the ventral-lateral regions of the embryo. (A, B) Patterning of the neuroectoderm along the AP and DV axes of the embryo occurs at early stage of embryogenesis. These cells are subdivided and positioned into neural equivalence group by a Cartesian coordinate system (A). In the VNC AP axis patterning, the gap and pair-rule genes define the segment-polarity genes expression within each segment, while the segment-polarity genes define the fate of neuronectodermal cells in each row. The combination of the nuclear factor NF-κB, BMP and the EGFR (A) divides neuroectoderm into three longitudinal stripes, ventral, intermediate and dorsal columns by controlling the expression of *vnd*, *ind* and *msh*, respectively (B). (C) The combination of two groups of genes, *ac/sc* Complex and Notch/Delta signaling pathway, leads to the selection of one cell within the proneural cluster developing as NB, and migrate to the interior of the embryos. (D) Five sequential waves of NB segregation can produce about 30 NBs in each hemisegment. (E) Each NB follows an invariant stem cell lineage and forms GMC in the embryo, which then produces a specific pair of postmitotic cells, neurons/glia. The specification of GMCs is controlled by five transcription factors, *hb*, *Kr*, *Pdm*, *cas* and *grh*, which are sequentially expressed in a temporal cascade in each NB, and then transmitted to the GMC and postmitotic cells. (F) The asymmetric division of GMC, under the control of the antagonistic interactions between the Notch pathway and Numb, specifies the different identities of glial and neuronal cells. In the end, there are about 30 glia and 350 neurons in each hemisegment of the VNC.

It has been reported that each thorax hemisegment generates about 100 more postmitotic cells than abdominal hemisegment (Skeath & Thor 2003). Although the ground state for each NB of different segments is the same, several NBs (NB1-1, NB2-2, NB6-4 and NB3-1) show segmental differences (Bossing et al. 1996). This segmental variety of lineages along VNC is under *Hox* gene control. For example, apart from generating the aCC motoneuron and pCC intersegmental interneuron, NB1-1 produces a cluster of local interneurons and the SPG-A and SPG-B glial cells in abdominal segments while only generates another motorneuron CoA in thorax (Schmid et al. 1999). The activity of *Ubx* and *abd-A* determine the composition difference of neuroblast NB1-1 between the thoracic and abdominal tagmata (Prokop & Technau 1994; Udolph et al. 1993). The thoracic and abdominal lineage of NB6-4 is specified by *abd-A* and *Abd-B* (Berger et al. 2005a; Berger et al. 2005b). In thorax, NB6-4 is the ground state, the expression of *CyclinE* (*CycE*) is necessary and sufficient to trigger the asymmetric division of NB6-4, and promote neuronal development. However, the *CycE* expression is blocked by *abd-A* and *Abd-B* in the abdominal segments, which leads to symmetric division of NB6-4 and produces two glial cells.

Hox genes also have been shown to control multiple processes of neural sepcification of the postmitotic cell (Miguel-Aliaga & Thor 2004; Rogulja-Ortmann et al. 2008; Suska et al. 2011). *Hox* genes are known to control neuronal apoptosis. It has been shown that *Ubx* and *Antp* act antogonistically to control the segmental specificity of NB7-3 and NB2-4 cell lineages development in the *Drosophila* embryonic CNS (Rogulja-Ortmann et al. 2008). *Ubx* is necessary to activate *reaper*-dependent cell death, while *Antp* controls the survival of these cells.

The aforementioned examples highlight the *Hox* genetic input in the control segment variety among VNC cell lineages during embryogenesis. Additionally, *Hox* genes are also involved in control of VNC segment-specificity during postembryonic development. For example, by manipulating *Ubx* expression in postembryonic VNC, Marin *et al.* found that *Ubx* determines the segment specific feature of neuron morphology and survival pattern (Marin et al. 2012).

Evidence provided above shows that *Hox* regulation plays important roles during CNS development. However, few studies address the molecular mechanisms underlying the modulation of *Hox* expression in the developing CNS. A study suggests that modulation of different length of *Ubx* 3' UTR providing accessibility of different sets of regulatory regions, thus controlling the *Ubx* expression during *Drosophila* CNS development (Thomsen et al. 2010). In the next section, I will introduce the function of 3' UTR in the gene regulation in detail.

1.6 3' untranslated regions (3' UTRs)

3' UTR of a gene contains regulatory information that determines mRNA fate. For instance, RNA binding protein (RBP) binding sites or miRNA binding sites that are involved in post-transcriptional regulation are often located within the 3' UTR (Alonso 2012; Moore 2005). Alternative polyadenylation (APA) is a mechanism that leads to the generation of alternative 3' UTRs for the same gene, providing an additional level of gene expression control within 3' UTRs. This mechanism is widespread and around half of human protein coding genes have more than one polyadenylation site (Giammartino et al. 2011). Genes that undergo APA express mRNAs with alternative 3' UTRs, carrying distinct combinations of post-transcriptional *cis*-regulatory elements (Majoros & Ohler 2007). For example, proliferating cells express mRNA with shortened 3' UTR, which have fewer miRNA target sites and thus decrease the impact of miRNA-mediated regulation (Sandberg et al. 2008). It has also been reported that miRNAs can shape the 3' UTR sequences of genes during the evolution. For instance, ubiquitously expressed genes, like housekeeping genes, tend to have shorter 3' UTRs that are depleted in miRNA target sites and thus avoid miRNA-based regulation, while miRNA-targeted developmentally regulated genes often have longer 3' UTRs (Stark et al. 2005). Therefore, APA of the 3' UTRs is greatly associated with miRNA-mediated post-transcriptional regulation.

It has been shown that the length of *Ubx* 3' UTR can be alternatively selected during the post-transcriptional process of APA (Kornfeld et al. 1989; O'Connor

et al. 1988). *Ubx* has two polyadenylation sites in the 3' UTR, which generate two *Ubx* 3' UTRs with different lengths, commonly named the *Ubx* short 3' UTR (1011bp) and *Ubx* long 3' UTR (2149bp), respectively. These two *Ubx* 3' UTRs contain two different sets of regulatory information that allow *Ubx* to be differentially regulated during animal development. A study has shown that different *Ubx* 3' UTR isoforms are differentially expressed during *Drosophila* embryonic CNS development (Thomsen et al. 2010). The *Ubx* short 3' UTR is expressed throughout embryonic development in different tissues, while the *Ubx* long 3' UTR is exclusively expressed in the CNS. Overexpression of a *Ubx* long 3' UTRs reporter construct in the *Drosophila* embryonic CNS has been shown a gradual decrease of expression of this construct towards the posterior of the embryo, which is consistent with the targeted repression by the the regulatory miRNA, *miR-iab-4* and *miR-iab-8*, which display posterior expression patterns.

Interestingly, another three *Hox* genes, *Antp*, *abd-A* and *Abd-B*, also display similar 3' UTR processing events in developing *Drosophila* embryos, indicating that this kind of 3' UTR processing may be a recurrent theme in *Hox* developmental regulation. Another study from the host lab reports that the *Ubx* polyadenylation signals that control APA in *cis* are conserved across 12 different *Drosophila* species, suggesting that this is an ancestral feature of *Drosophila Hox* gene regulation (Patraquim et al. 2011).

The alternative processing 3' UTR of a gene leads to the presence or absence of the potential miRNA target sites to modulate interactions between particular miRNA species and subsets of mRNAs isoforms (Bartel & Chen 2004). However, the question of how miRNAs find their target sites on the 3' UTRs of mRNAs is still unresolved. In the next sections, I am going to introduce the miRNAs and currect bioinformatic programs designed to find the miRNAs that target the 3' UTR of a particular mRNA.

1.7 miRNAs

The groundbreaking discovery of the first miRNA in 1993 reported that the small

regulatory RNA *lin-4* acts on the *lin-14* 3' UTR in a complementary manner, controlling the timing of larval development in *Caenorhabditis elegans* (*C.elegans*) (Lee 1993; Wightman et al. 1993). Within seven years of the discovery of *lin-4*, no evidence for the existence of miRNAs similar to *lin-4* was collected, neither within nor beyond nematodes (Bartel 2004). Upon the discovery that a novel miRNA, *let-7*, controlls the timing of *C. elegans* development in 2000 (Reinhart et al. 2000; Slack et al. 2000), homologs of *let-7* were detected in a wide range of bilateral animals with conserved temporal regulation, suggesting that *let-7* may control late temporal transitions across the animal phylogeny (Pasquinelli et al. 2000). One year later, a large number of these small endogenous RNAs, termed miRNAs, were detected in invertebrates and vertebrates (Lagos-Quintana et al. 2001; Lau et al. 2001; Lee & Ambros 2001), implying that miRNAs may have broad regulatory functions in animals.

miRNAs are endogenous small noncoding RNAs, approximately 23 nucleotides long, which act as post-transcriptional regulators of gene expression. The biogenesis of metazoan miRNAs is a complex process (Kim et al. 2009; Krol et al. 2010; Ameres & Zamore 2013). Primary miRNAs (pri-miRNAs) are transcribed from independent genes by RNA polymerase II (RNAP II) (Lee, Y. et al. 2004) and processed to a hairpin-shaped ~70bp precursor miRNAs (pre-miRNAs) by the Drosha-DGCR8 complex (Drosha-Pasha complex in *Drosophila*) (Han et al. 2004). Then the pre-miRNAs are recognized by the Exportin-5 and exported from the nucleus to cytoplasm (Kim 2004). The nuclear processing for intronic miRNAs (from the introns of protein-coding genes) or miRtrons, is slight different, which undergo splicing and debranching to form the pre-miRNAs instead of being processed by Drosha-DGCR8 complex (Figure 1.6A). The cleavages of pre-miRNAs are mediated by Dicer (Dicer-1 in *Drosophila*), Argonaute 1-4 (AGO 1-4) (AGO1 in *Drosophila*) and TAR RNA-binding protein (TRBP) or PACT (LOQS in *Drosophila*) in the cytoplasm to yield the miRNA/miRNA* duplexes (Figure 1.6B) (Chendrimada et al. 2005). Being loaded to the RNA-induced silencing complex (RISC), miRNAs stimulate the mRNAs degradation or repress translation, while miRNAs* are either degraded or loaded into the RISC, as detailed above (Okamura et al. 2009; Okamura et al. 2008). Functional mature miRNAs are processed from one or both arms of the

pre-miRNA hairpin, thus, there are different combinations of functional mature miRNAs, including 3' arm, 5' arm or both 3' and 5' arm depending on the maturation of the pre-miRNA (Griffiths-Jones et al. 2011; Ameres & Zamore 2013).

By binding to the target mRNA in a sequence-complementary manner, miRNAs manipulate mRNA expression levels by either repressing mRNA translation or mediating mRNA degradation (Bartel 2009; Filipowicz et al. 2008). It is speculated that miRNAs probably regulate approximately 30 percent of all protein-coding genes, playing major roles in the post-transcriptional regulation of many biological processes (Nilsen 2007; Filipowicz et al. 2008).

The number of miRNAs in an individual organism is positively correlated with its morphological complexity, which suggests that the establishment of the complex gene regulatory networks of multicellular organisms requires the post-transcriptional regulation provided by miRNAs (Kosik 2009). miRNAs comprise about 1% of all animal genes and mutations of proteins required for the miRNA biogenesis or function lead to defects in animal development (Okamura et al. 2004; Liu et al. 2004; Wienholds et al. 2003; Grishok et al. 2001; Lee, Y. S. et al. 2004; Bernstein et al. 2003), highlighting the role of the miRNA gene-class during animal development. It has been shown that miRNAs participate in multiple biological processes of different cellular and developmental contexts, such as developmental timing, cell death and proliferation, morphogenesis and cell differentiation (Ambros 2004; Alvarez-Garcia & Miska 2005). For example, in *Drosophila*, *bantam* acts as an anti-apoptotic miRNA by targeting the pro-apoptotic *head involution defective* (*hid)* transcripts. *bantam* both enhances cell proliferation and represses apoptosis to promote tissue growth (Brennecke et al. 2003). In addition, *miR-34* is required in the regulation of senescence in general and long-term brain integrity in particular, which provides a molecular link between aging and neurodegeneration. Indeed, *miR-34* loss of function prompts accelerated brain aging, neurodegeneration and decline in survival (Liu et al. 2012).

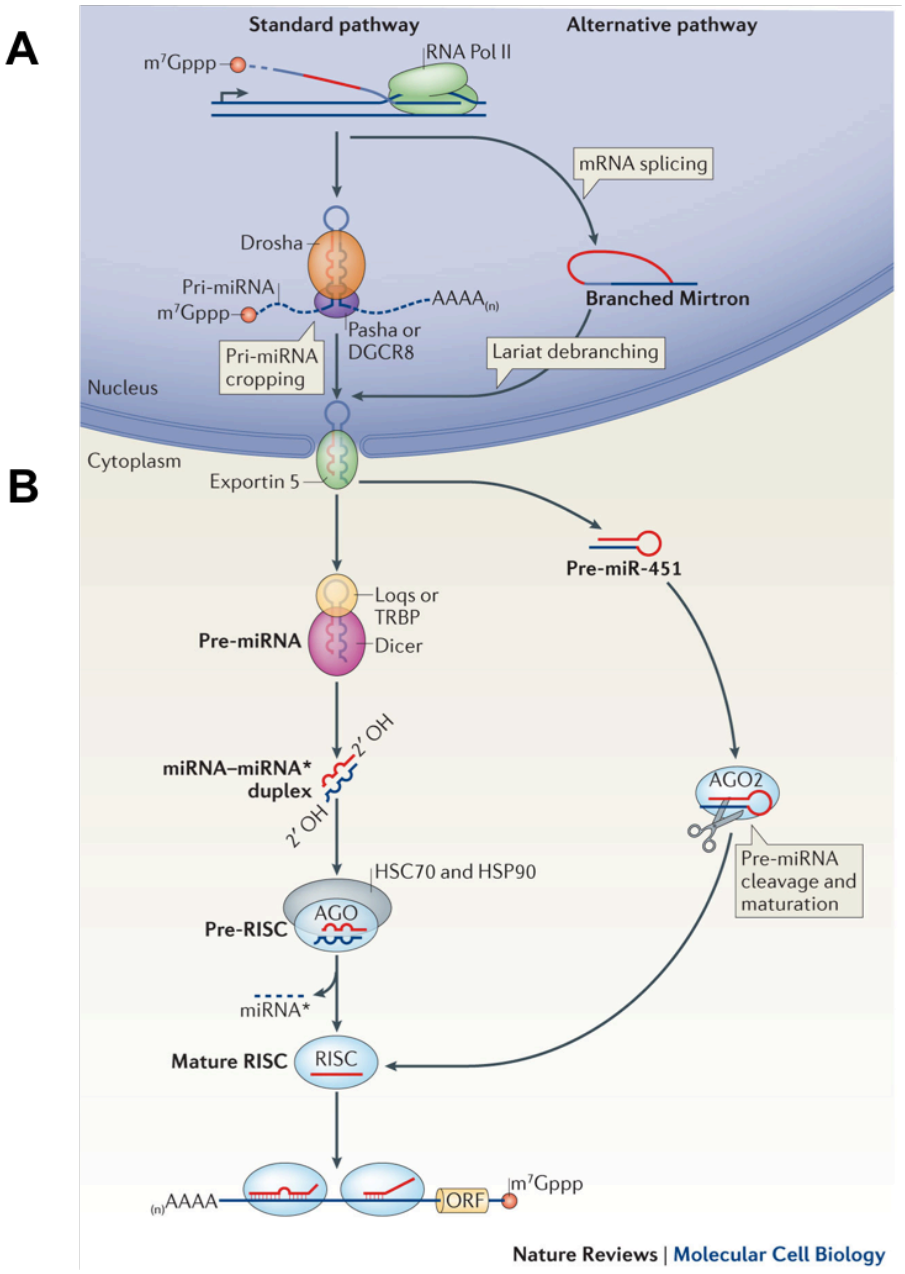


Figure 1.6 The biogenesis of miRNAs.

Taken from Ameres & Zamore (Ameres & Zamore 2013). (A) pri-miRNAs are transcribed from independent genes by RNAP II or from the introns of protein-coding genes. In the canonical processing pathway, the pri-miRNAs are transcribed from the genome by RNAP II and processed to a hairpin-shaped ~70bp pre-miRNAs by the Drosha-DGCR8 complex (Drosha-Pasha complex in *Drosophila*). The pre-miRNAs are recognized by the Exportin-5 and exported from the nucleus to cytoplasm. The nuclear processing for intronic miRNAs, or miRtrons, is slight different. Instead of being processed by Drosha-DGCR8 complex, the transcribed intronic miRNAs undergo splicing and debranching to form the pre-miRNAs. (B) In the cytoplasm, Dicer (Dicer-1 in *Drosophila*), AGO 1-4 (AGO1 in *Drosophila*) and TRBP or PACT (LOQS in *Drosophila*) mediate the cleavage of pre-miRNAs to yield the miRNA/miRNA* duplexes. miRNAs are loaded to the RISC and mediate the post-transcriptional regulation of target mRNAs, while miRNAs* are either degraded or loaded into the RISC, as detailed above.

Conversely, a current study suggests that miRNAs could upregulate or derepress gene expression (Vasudevan 2012). For example, *miR-206* binds the 3' UTR of *KLF4* mRNA and upregulates its translation in quiescent mammalian cell lines (Lin et al. 2011). Therefore, it is possible that miRNAs can both activate or repress gene expression compared to the previous unique repression hypothesis (Bartel 2009).

However, the detailed mechanism of regulation achieved by miRNAs in gene expression is still not fully understood. Some published reports have contradictory results in terms of the effects of miRNAs on mRNA translation and stability (Filipowicz et al. 2008). Therefore, a comprehensive and consensual mechanism of miRNAs regulation has still to be established, especially how miRNAs recognize their targets. In the next sections, I am going to use bioinformatic approaches to explore this question.

1.8 Available miRNA target prediction programs

In miRNA biology, miRNA function depends on the mRNA repertoire regulated. Therefore, it is essential to define the rules how miRNAs recognize their targets. For this reason, bioinformatic approaches have been developed to tackle these rules. In this section, I have reviewed all currently available softwares for miRNA target prediction. All miRNA target prediction programs that are still commonly used are listed in Table 1.1. Most programs, such as ComiR, DIANA, Microinspector, microTAR, miRanda, miRDB, miRecord, miRNAMap, miRtrail, miRWalk, PicTar, PITA, RNA22, RNAhybrid, TargetScan and STarMir, could be used for the prediction of miRNA targets in various species, assuming that the mechanism for miRNA targeting is conserved across species. Understanding the mechanism for one species could potentially lead to a better, general understanding miRNA targeting for other species as well.

Table 1.1 Current target prediction programs available for public access.

Program	Organism[a]	website	reference
ComiR[b]	h, m, f, w	http://www.benoslab.pitt.edu/comir/index2.php	(Coronnello & Benos 2013; Coronnello et al. 2012)
DIANA	h, m, w, f	http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=microT_CDS/index	(Reczko et al. 2012; Paraskevopoulou et al. 2013)
EMBL	f	http://www.russelllab.org/miRNAs/	(Brennecke et al. 2005)
GenMiR + +	h	http://www.psi.toronto.edu/genmir/	(Huang et al. 2007)
HOCTAR	h	http://hoctar.tigem.it/	(Gennarino et al. 2011)
Magia	h	http://gencomp.bio.unipd.it/magia/start/	(Sales et al. 2010)
Microinspector	any	http://bioinfo.uni-plovdiv.bg/microinspector/	(Rusinov et al. 2005)
microTAR	m, f, w	http://tiger.dbs.nus.edu.sg/microtar/	(Thadani & Tammi 2006)
miRanda	h, m, r, f, w	http://www.microrna.org/microrna/home.do	(Betel et al. 2010; Betel et al. 2008; Enright et al. 2003; John et al. 2004)
miRDB	h, m, d, r, c	http://mirdb.org/miRDB/	(Wong & Wang 2014)
miRecord[b]	h, m, f, r, z, w, c, s, d	http://c1.accurascience.com/miRecords/	(Xiao et al. 2009)
miRGator	h, m	http://genome.ewha.ac.kr/miRGator/miRNAexpression.html	(Nam et al. 2008)
miRGen	h, m	http://www.microrna.gr/mirgen/	(Alexiou et al. 2009)
miRNAMap 2.0	any	http://mirnamap.mbc.nctu.edu.tw/index.php	(Hsu et al. 2008)
miRTar	h	http://mirtar.mbc.nctu.edu.tw/human/index.php	(Fontana et al. 2007)
miRtrail	h, m, z	http://mirtrail.bioinf.uni-sb.de/mirtrail.php	(Backes et al. 2007)
miRWalk2.0[b]	h, m, r	http://www.umm.uni-heidelberg.de/apps/zmf/mirwalk/index.html	(Dweep et al. 2011)
mirWIP	w	http://146.189.76.171/query.php	(Hammell et al. 2008)
PicTar	f, v, w	http://pictar.mdc-berlin.de/	(Grün et al. 2005; Krek et al. 2005)
PITA	h, m, f, w	http://genie.weizmann.ac.il/pubs/mir07/mir07_prediction.	(Kertesz et al. 2007)

		html	
RNA22	h, m, f, w	https://cm.jefferson.edu/rna22/	(Miranda et al. 2006)
RNAhybrid	h, f, w	http://bibiserv.techfak.uni-bielefeld.de/rnahybrid?id=rnahybrid_view_download	(Rehmsmeier et al. 2004)
STarMir	w, h, m	http://sfold.wadsworth.org/cgi-bin/starmir.pl	(Rennie et al. 2014)
TagetScan	m, w, f, z	http://www.targetscan.org/fly_12/	(Lewis et al. 2005; Kheradpour et al. 2007)
TargetMiner	h,	http://www.isical.ac.in/~bioinfo_miu/targetminer20.htm	(Bandyopadhyay & Mitra 2009)
TargetRank	h, m	http://genes.mit.edu/targetrank/	(Nielsen et al. 2007)

Programs are listed in alphabetical order by program name.

a. Organisms for which the program is applicable (h: human; m: mouse; d: dog; r: rat; c: chicken; f: fly; w: worm; z: zebrafish; v: vertebrate; s: sheep).

b. Programs that were developed with a combination of several prediction programs.

1.9 Description of the available miRNA target prediction programs for *Drosophila melanogaster*

The programs for the miRNA target prediction in *Drosophila melanogaster* are listed in Table 1.2. All ten programs have online tools, which provide a simple and quick prediction approach to evaluate the potential targets of a particular miRNA or potential miRNAs that may regulate a specific mRNA. miRanda and TargetScan provide the ability for the user to input miRNA sequences, which allows us to restrict our predictions to the miRNAs of interest. On the website of PITA, RNA22, RNAhybrid and ComiR, both miRNA input and targeted mRNA input options are available, which are very useful to analyse a small group miRNAs and mRNAs of interest. Furthermore, the developers of miRNA target prediction programs such as miRanda, PITA, RNA22 and RNAhybrid shared the executable resources that are convenient for users to download and predict miRNA targets locally, offering the ability to further customize the analysis. This function is especially useful for researchers who work on non-model species or generate novel miRNA annotations.

Most miRNA prediction tools focus on the 3’ UTR of the genes, except DIANA,

RNAhybrid and RNA22, which also consider the coding sequence (CDS) and CDS plus 5' untranslated region (5' UTR), respectively. Recently, the growth of high-throughput approaches (Hafner et al. 2010; Forman & Coller 2010; Wang et al. 2010) and individual miRNA case studies (Abdelmohsen et al. 2010; Elcheva et al. 2009; Tay et al. 2008; Zhou & Rigoutsos 2014; Ørom et al. 2008), led to an accumulation of experimental evidence suggesting that apart from target 3' UTRs, miRNAs may also regulate mRNA expression by targeting CDS and 5' UTR. Hafner et al. reported that miRNAs binding sites occur on the 3' UTR and CDS of target mRNAs in almost similar proportions (Hafner et al. 2010). Also, Clark and colleagues (Clark et al. 2014) reported that 50% of miRNA target sites are located in the CDSs and 5' UTRs. Taking the CDS and 5' UTR of mRNAs into account for miRNA target prediction is therefore crucial to understand the mechanism of miRNA targeting. The detailed features of each miRNA program are summarized below (Table 1.2).

Table 1.2 Resource availability for miRNA target prediction programs used in this study

Program	Version	predicted region	Web access			download tool
			online tool	miRNA input	mRNA input	
DIANA	DIANA-microT web server v5.0	3' UTR and CDS	Y	N	N	N
EMBL	updated in 2005	3' UTR	Y	N	N	N
miRanda	miRanda-3.3a	3' UTR	Y	Y	N	Y
PicTar	updated on March 26, 2007	3' UTR	Y	N	N	N
PITA	self-input web prediction	3' UTR	Y	Y	Y	Y
RNA22	remoteRNA22v2	3' UTR, 5' UTR and CDS	Y	Y[a]	Y	Y
TargetScan	TargetScanFly Release 6.2	3' UTR	Y	Y[b]	N	N
RNAhybrid	RNAhybrid Windows (32 Bit) binary package	3' UTR, 5' UTR and CDS	Y	Y	Y	Y
ComiR	updated on 21 July, 2014	NA	Y	Y	Y	N
miRecord	updated on April 27, 2013	NA	Y	N	N	N

'Y' equates to yes, while 'N' means No. 'NA' indicates this category is not applicable in this program. (a. only allows to input up to 50 miRNAs; b. only input one miRNA per time.)

1.10 Comparison of the features of different bioinformatics prediction tools (the detailed features of each programs see the method)

Understanding the mechanism of miRNA target site recognition has been a big challenge. One reason for this difficulty is the limited sequence complementarity between miRNA and mRNA that mediates the interaction of miRNA and its targets. Therefore, it is possible that there are hundreds of potential target genes per miRNA and a specific gene may also be predicted to be targeted by a large number of miRNAs. In fact, a number of bioinformatic and experimental methods have been developed to address how miRNAs identify their targets. To have a overview of important parameters for detecting miRNA targets, I divided the miRNA targeting features of ten bioinformatics programs that are available for miRNA prediciton of *Drosophila* mRNA targets into four categories, based on the main parameters used in miRNA target predictions: sequence, thermodynamics, conservation and others (Table 1.3).

1.10.1 Sequence

I further divided this category into five subcategories, including sequence complementary, AU flanking, target motif, target site position and multiple sites.

Sequence complementary is the most essential feature for target recognition. Using this feature, miRNAs target prediction tools consider the Watson-Crick pairing of the 5’ region of mature miRNAs to mRNA target sequences, centered on the 2-7 nucleotide of miRNA that is defined as miRNA seed (Brennecke et al. 2005). With the exception of miRecord, which concatenates the target predictions of 11 different miRNA target prediction algorithms, the remaining 9 programs all take this 5’ dominant sequence complementary feature into account. In addition to this stringent-seed pairing, some miRNA target prediction tools also consider the moderate-stringent-seed paring, such as one G:U, one bulge or one mismatch in the seed region when there are 7-8 seed pairings. DIANNA, EMBL and miRanda tolerate any of these three, while the PicTar only allows one bulge or one mismatch, PITA only accepts one G:U and RNAhybrid

only tolerates one bulge. The algorithm design for RNA22 is an exception, which not only tolerates any situation such as one G:U, one bulge or one mismatch in the seed region, additionally, it also allows for more than one G:U wobbles in the seed region.

Grimson *et al.* reported that additional Watson-Crick pairing between the mRNA and the 3' part of miRNA greatly enhance the efficacy of miRNA target recognition in addition to seed pairing (Grimson et al. 2007). Thus, 3' compensatory sites become another type of sequence complementary mechanism that do not require stringent 5' pairing, but need perfect 3' matching. Programs based on this feature include DIANA, EMBL, miRanda and RNA22.

Lewis *et al.* found that AU sequence flanking the seed complementary sequence is conserved across five vertebrate species (human, mouse, rat, dog and chicken), indicating that this feature could be important for miRNA target sites recognition (Lewis et al. 2005). miRNA targeting prediction programs such as DIANA, miRanda, TargetScan and ComiR all include this feature in their prediction. Some programs also use the identity of extended flanking sequences for miRNA targeting prediction, e.g. DIANA and TargetScan. As miRNAs bind mRNAs through RISC, the unpairing of the flanking sequence around the seed region, which allows RISC access to the target site, is another feature to be taken into account. For instance, PITA offers the option to calculate the energetic cost of unpairing the flanking sequence, which enhances the accuracy of this software's prediction.

Another sequence identity feature, the miRNA target motif, is also important to consider for miRNA target predictions. TargetScan identifies the general target motif in the 3' UTR within which miRNA target sites may be found, using phylogenetic conservation features. DIANA uses the PAR-CLIP data to find the true miRNA-recognition elements (MREs). RNA 22, instead of using the currently available miRNA identity, analyses the feature of miRNA sequence and matches this character to the 3' UTR to find potential miRNA target sites, even those target sites for miRNAs are not yet discovered.

Gaidatzis *et al.* reported that miRNA target sites have the propensity to locate close to the start and end of the 3' UTR (Gaidatzis et al. 2007). DIANA, miRanda, TargetScan and ComiR incorporate this feature in their miRNA target predictions. Not only the position of target sites, but also the distance between target sites affects targeting efficacy. Sætrom *et al.* found that a distance among two seed sites between 13 and 35 nucleotides could optimize the target down-regulation (Sætrom et al. 2007). From all ten programs, only DIANA takes this feature into consideration.

This kind of synergistic miRNA effect does not only apply to multiple target sites of a single miRNA, but is also suitable to predict different miRNAs that may regulate the same mRNA. One possible mechanism for optimal miRNA-mediated down regulation is that multiple RISC binding 3' UTR targets in close distance could cooperatively support each other and therefore enhance regulatory strength (Saito & Sætrom 2010). miRanda, PicTar and ComiR incorporate this feature into their respective miRNA target predictions.

1.10.2 Thermodynamics

Target site accessibility is another feature that should be taken into consideration in miRNA target site predictions. In addition to seed complementary, the mRNA structural unpairing in order for the miRNA to gain access to a specific mRNA target is energetically costly. 5' seed sequence is crucial for miRNA-mRNA interaction (Brennecke et al. 2005). Kertesz *et al.* reported that the effect of a single nucleotide deletion, insertion or mutation is similar to that of the closing structure, which indicates that target site accessibility is very important for miRNA targeted gene repression and maybe as important as seed complementarity (Kertesz et al. 2007). PITA, one of the earliest programs to take target site accessibility into account, calculates the energy loss of unpairing mRNA secondary structure and compares it with the energetic gain from a successful miRNA-mRNA duplex formation. Taking advantage of the same algorithm from PITA, ComiR also utilizes this feature in miRNA target predictions.

1.10.3 Evolutionary conservation

It was reported (Kheradpour et al. 2007) that the functional elements are usually conserved during evolution, so the first few developed miRNA prediction programs such as EMBL, TargetScan, DIANA and PicTac rely heavily on the evolutionary conservation of miRNA target sites for prediction. However, this filter only applies to common target sites among species, and those species-specific miRNA target sites were ignored completely. Therefore, miRanda, PITA, RNA22 and ComiR, developed a new algorithm that can find other target sites in addition to the consideration of target sequence conservation. This new way of miRNA target prediction greatly increased the number of target sites and discovered huge amount of non-canonical and non-conserved target sites, significantly increased the chance of discovering genuine miRNA and mRNA interactions.

1.10.4 Unique features of each program

Other than the previously described features, some programs developed their unique features to improve the prediction efficiency and accuracy. For instance, DIANA combines the score of individual prediction from 3' UTR and CDS, which considers the miRNA synergistic effect to improve the accuracy of prediction. ComiR incorporates the miRNA expression level in the prediction. This feature takes the expression level of miRNA into account and could monitor miRNA regulation in a cellular and tissue specific level. Another feature of ComiR is that this program includes support vector machine (SVM), a machine learning method, which can recognize the pattern by the feature of multiple dimensions and find the maximum-margin hyperplane to get a better classification. Support vector regression (SVR) is a version of SVM for regression. By training the data from the mRNA expression level change following the miRNA transfection, miRanda and ComiR integrated mirSVR with the other features for miRNA prediction. With a combination of features of multiple miRNA targeting programs so far, these two miRNA prediction tools are most likely to have good performance by taking advantage of the advanced algorithms.

miRNAs have a significant impact on 3’ UTR evolution. miRNA target genes usually have longer 3’ UTR with targeting sites, while house-keeping genes tend to have shorter 3’ UTR (Stark et al. 2005). Thus, it might be beneficial to taking into account the length of 3’ UTR when performing miRNA predictions. miRanda, RNA22, RNAhybrid and ComiR programs put this feature into developing the prediction algorithm. In the prediction of miRanda, the program predicts the thermodynamically most favorable hybridisation sites by calculating the minimum free energy (MFE). The longer 3’ UTRs are tending to contain much more good MFE binding sites, so miRanda considers the 3’ UTR length for miRNA target prediction to avoid the noise of different 3’ UTR lengths. For the case of RNA 22, by using the miRNA sequence feature to find the target site, the longer the 3’ UTR, the more likely that I can find spurious binding sites. Therefore, the RNA 22 developer takes the length of 3’ UTR into account for miRNA target prediction and applies this feature to predict 5’ UTR and CDS regions as well.

The purpose for development of miRNA target program is, by understanding the mechanism of miRNA recognition, to find the genuine target site and apply this miRNA and mRNA interaction to study the endogenous miRNA regulation. Therefore, the *in vivo* validation of the algorithm is crucial to evaluate the performance of the program. miRanda, PicTar, PITA and TargetScan programs use *in vivo* evidence to improve the efficiency of the program. For the case of EMBL, the developer uses *in vivo* experimental data to improve the performance of the program.

However, due to the lack of high-throughput experimental data, the update of miRNA prediction alghorithms is dependent on case studies, significantly dampening prediction accuracy (Sethupathy et al. 2006). Experimental approaches may be a much better option to study the regulatory miRNAs for a particular gene.

Table 1.3 Summary of the features used by the *Drosophila* target prediction programs.

miRNA target prediction programs	DIANA	EMBL	miRanda	PicTar	PITA	RNA22	targetScan	RNAhybrid	ComiR
Sequence									
Sequence complementarity									
5' dominant (canonical and seed)	Y	Y	Y	Y	Y	Y	Y	Y	Y
3' compensatory	Y	Y	Y	N	N	Y	N	N	N
7-8 seed pairing with 1G:U, 1 bugle or 1 mismatch	Y	Y	Y	Y[d]	Y[e]	Y[f]	N	Y[l]	N
Flanking									
AU flanking	Y[a]	N	Y	N	N	N	Y	N	Y
flanking sequence	Y	N	N	N	Y[j]	N	Y	N	Y
Target motif									
target motif	Y[a]	N	N	N	N	Y[k]	Y	N	Y
Target position									
relative position between MRE	Y[a]	N	N	N	N	N	N	N	N
relative distance of target position from end of 3' UTR	Y[a]	N	Y	N	N	N	Y	N	Y
Cooperative regulation									
cooperative regulation	N	N	Y	Y	N	N	N	N	Y
Thermodynamics									
ΔΔG	N	N	N	N	Y	N	N	N	Y
Conservation									
target site conservation	Y[a]	Y	Y[b]	Y	N	N	Y[h]	N	N
non-canonical and non-conserved target site	N	N	Y	N	Y[g]	Y	N	N	Y

Table 1.3 Summary of the features used by the *Drosophila* target prediction programs.

Others

combine the score of 3' UTR and CDS	Y	N	N	N	N	N	N	N	N
Ranking by down-regulation score (SVR)	N	N	Y	N	N	N	N	N	Y
miRNA expression level	N	N	N	N	N	N	N	N	Y
in vivo evidence	N	Y	Y[c]	Y[c]	Y[c]	N	Y[c]	N	N
multiple prediction algorithm	N	N	Y	N	N	N	N	N	Y
support vector machine (SVM)	N	N	N	N	N	N	N	N	Y
length of 3' UTR	N	N	Y	N	N	Y	N	Y	Y

The prediction parameters listed above are used for all the miRNA prediction programs in this study except for miRecord, which integrates the predicted targets of the following miRNA target prediction tools: DIANA-microT, MicroInspector, miRanda, MirTarget2, miTarget, NBmiRTar, PicTar, PITA, RNA22, RNAhybrid, and TargetScan/TargertScanS. 'Y' indicates that the feature is used in the program, while 'N' means the feature is not considered. (a. The feature is considered for both the 3' UTR and CDS; b. the conservation of target sites is a feature rather than a filter; c. these programs use *in vivo* experimental data as a validation rather than an approach to develop the program; d. no tolerance for the G:U in the seed region; e. only tolerance for G:U in the seed region; f. no restriction of G:U number in the seed region; g. only consider the non-conserved sites; h. consider the conserved target motif; i. only tolerance for one bulge in the seed region; j. consider the energy costs for the unpairing of the flanking sequence; k. use the miRNA features to identify the target islands in the mRNAs. The miRNA sequence features identify target island, while miRNA target motif is pairing to the known miRNA sequences).

1.11 Aims of this study

As I have described above, miRNA-mediated post-transcriptional regulation is a crucial layer in the establishment of gene expression patterns during animal development. However, the mechanisms that allow miRNAs to identify their targets are still unclear. This makes it difficult to determine the full repertoire of active miRNAs controlling gene expression profiles within the dynamic context of development. Here I approach this question focusing on the identification of miRNAs able to control the expression of the *Drosophila Hox* gene, *Ubx,* during development.

In Chapter 3 and Chapter 4, I ask: (i) Can I identify novel *Ubx*-regulatory miRNAs using a genetic screen? (ii) Of those candidate miRNAs identified in our screen, are they able to affect *Ubx* expression? (iii) Do candidate *Ubx*-regulatory miRNAs interact with *Ubx* 3' UTR directly?

In the fifth chapter, I investigate the effects of two *Ubx*-regulatory miRNAs with strong regulatory effects in the genetic screen (Chapter 3), elucidating their roles in the establishment of *Hox* expression patterns during the embryonic development of the *Drosophila* CNS.

In addition, I evaluate different miRNA target-prediction methods by comparing *state-of-the-art* bioinformatics platforms with the outcome of our genetic screen.

Overall, my work makes a contribution to the understanding of miRNA-target interactions within the physiological context of development.

CHAPTER 2 General methods:

2.1 Stocks and Fly Husbandry

Fruit flies (*Drosophila melanogaster*) were cultured using molasses fly food following the standard protocols at 25°C on a 12 hours light and dark cycles. Most of the UAS-miRNA lines used for the genetic screen were obtained from the Bloomington Stock Center (http://flystocks.bio.indiana.edu/Browse/uas/uas_mir.php) (Bejarano et al. 2012). Oregon Red (Or-R) strain was used as a control. All fly stocks used were listed in Table 2.1.

Table 2.1 Fly stocks used in this thesis.

ID	Genotype	Origin
UAS-bft	w [1118]; + ; UAS-bft	Bloomington # 41133
UAS-let-7	w [1118]; + ; UAS-let-7	Bloomington # 41171
UAS-miR-1	w [1118]; + ; UAS-miR-1/TM3, Sb[1]	Bloomington # 41125
UAS-miR-10	w [1118]; + ; UAS-miR-10/TM3, Sb[1]	Bloomington # 41169
UAS-miR-100	w [1118]; + ; UAS-miR-100	Bloomington # 41166
UAS-miR-1000	w [1118]; + ; UAS-miR-1000	Bloomington # 41201
UAS-miR-1001	w [1118]; + ; UAS-miR-1001	Bloomington # 41202
UAS-miR-1003	w [1118]; + ; UAS-miR-1003	Bloomington # 41220
UAS-miR-1004	w [1118]; + ; UAS-miR-1004	Bloomington # 41203
UAS-miR-1006	w [1118]; + ; UAS-miR-1006	Bloomington # 41204
UAS-miR-1007	w [1118]; + ; UAS-miR-1007	Bloomington # 41221
UAS-miR-1009	w [1118]; + ; UAS-miR-1009	Bloomington # 41205
UAS-miR-1010	w [1118]; + ; UAS-miR-1010	Bloomington # 41206
UAS-miR-1011	w [1118]; + ; UAS-miR-1011	Bloomington # 41210
UAS-miR-1012	w [1118]; + ; UAS-miR-1012	Bloomington # 41214
UAS-miR-1013	w [1118]; + ; UAS-miR-1013	Bloomington # 41215
UAS-miR-1015	w [1118]; + ; UAS-miR-1015	Bloomington # 41207
UAS-miR-1017	w [1118]; + ; UAS-miR-1017/TM3, Sb[1]	Bloomington # 41208
UAS-miR-12	w [1118]; + ; UAS-miR-12	Bloomington # 41140
UAS-miR-124	w [1118]; + ; UAS-miR-124	Bloomington # 41126
UAS-miR-133	w [1118]; + ; UAS-miR-133	Bloomington # 41132
UAS-miR-137	w [1118]; + ; UAS-miR-137	Bloomington # 41209
UAS-miR-14	w [1118]; + ; UAS-miR-14	Bloomington # 41178
UAS-miR-184	w [1118]; + ; UAS-miR-184	Bloomington # 41174

UAS-miR-190	w [1118]; + ; UAS-miR-190	Bloomington # 41154
UAS-miR-210	w [1118]; + ; UAS-miR-210	Bloomington # 41179
UAS-miR-252	w [1118]; + ; UAS-miR-252	Bloomington # 41127
UAS-miR-263b	w [1118]; + ; UAS-miR-263b	Bloomington # 41146
UAS-miR-274	w [1118]; + ; UAS-miR-274	Bloomington # 41172
UAS-miR-275, 305	w [1118]; + ; UAS-miR-275, 305	Bloomington # 41168
UAS-miR-275	w [1118]; + ; UAS-miR-275	Bloomington # 41142
UAS-miR-276a	w [1118]; + ; UAS-miR-276a	Bloomington # 41143
UAS-miR-276b	w [1118]; + ; UAS-miR-276b	Bloomington # 41162
UAS-miR-278	w [1118]; + ; UAS-miR-278	Bloomington # 41180
UAS-miR-279	w [1118]; + ; UAS-miR-279	Bloomington # 41147
UAS-miR-280	w [1118]; + ; UAS-miR-280	Bloomington # 41164
UAS-miR-281-1, 281-2	w [1118]; + ; UAS-miR-281-1, 281-2	Bloomington # 41167
UAS-miR-281-1	w [1118]; + ; UAS-miR-281-1	Bloomington # 41177
UAS-miR-282	w [1118]; + ; UAS-miR-282	Bloomington # 41165
UAS-miR-284	w [1118]; + ; UAS-miR-284	Bloomington # 41134
UAS-miR-285	w [1118]; + ; UAS-miR-285	Bloomington # 41160
UAS-miR-286	w [1118]; + ; UAS-miR-286	Bloomington # 41151
UAS-miR-2b-1	w [1118]; + ; UAS-miR-2b-1	Bloomington # 41128
UAS-miR-303	w [1118]; + ; UAS-miR-303	Bloomington # 41141
UAS-miR-303, 982	w [1118]; + ; UAS-miR-303, 982	Bloomington # 41193
UAS-miR-304	w [1118]; + ; UAS-miR-304	Bloomington # 41170
UAS-miR-305	w [1118]; + ; UAS-miR-305	Bloomington # 41152
UAS-miR-306, 9B, 9C, 79	w [1118]; + ; UAS-miR-306, 9b, 9c, 79	Bloomington # 41156
UAS-miR-307a	w [1118]; + ; UAS-miR-307a	Bloomington # 42026
UAS-miR-308	w [1118]; + ; UAS-miR-308	Bloomington # 41159
UAS-miR-308	w [1118]; + ; UAS-miR-308	Bloomington # 41809
UAS-miR-309	w [1118]; + ; UAS-miR-309	Bloomington # 41181
UAS-miR-310, 311, 312, 313	w [1118]; + ; UAS-miR-310, 311, 312, 313/TM3, Sb[1]	Bloomington # 41135
UAS-miR-310	w [1118]; + ; UAS-miR-310	Bloomington # 41155
UAS-miR-311	w [1118]; + ; UAS-miR-311	Bloomington # 41163
UAS-miR-312	w [1118]; + ; UAS-miR-312	Bloomington # 41144
UAS-miR-318	w [1118]; + ; UAS-miR-318/TM3, Sb[1]	Bloomington # 41161
UAS-miR-31a	w [1118]; + ; UAS-miR-31a	Bloomington # 42027
UAS-miR-31b	w [1118]; + ; UAS-miR-31b	Bloomington # 41129
UAS-miR-33	w [1118]; + ; UAS-miR-33	Bloomington # 41150
UAS-miR-34	w [1118]; + ; UAS-miR-34	Bloomington # 41158
UAS-miR-375	w [1118]; + ; UAS-miR-375	Bloomington # 41182
UAS-miR-4966, 975, 976, 977	w [1118]; + ; UAS-miR-4966, 975, 976, 977	Bloomington # 41223
UAS-miR-79	w [1118]; + ; UAS-miR-79	Bloomington # 41145
UAS-miR-8	w [1118]; + ; UAS-miR-8	Bloomington # 41176
UAS-miR-927	w [1118]; + ; UAS-miR-927	Bloomington # 41183

UAS-miR-929	w [1118]; + ; UAS-miR-929/TM3, Sb[1]	Bloomington # 41184
UAS-miR-92a	w [1118]; + ; UAS-miR-92a	Bloomington # 41153
UAS-miR-92b	w [1118]; + ; UAS-miR-92b	Bloomington # 41175
UAS-miR-932	w [1118]; + ; UAS-miR-932	Bloomington # 41157
UAS-miR-954	w [1118]; + ; UAS-miR-954	Bloomington # 41185
UAS-miR-955	w [1118]; + ; UAS-miR-955	Bloomington # 41186
UAS-miR-956	w [1118]; + ; UAS-miR-956	Bloomington # 41187
UAS-miR-958	w [1118]; + ; UAS-miR-958	Bloomington # 41222
UAS-miR-961	w [1118]; + ; UAS-miR-961	Bloomington # 41188
UAS-miR-963, 964	w [1118]; + ; UAS-miR-963, 964	Bloomington # 41216
UAS-miR-964	w [1118]; + ; UAS-miR-964	Bloomington # 41148
UAS-miR-966	w [1118]; + ; UAS-miR-966	Bloomington # 41211
UAS-miR-970	w [1118]; + ; UAS-miR-970	Bloomington # 41189
UAS-miR-973	w [1118]; + ; UAS-miR-973	Bloomington # 41190
UAS-miR-974	w [1118]; + ; UAS-miR-974	Bloomington # 41225
UAS-miR-976	w [1118]; + ; UAS-miR-976	Bloomington # 41149
UAS-miR-978	w [1118]; + ; UAS-miR-978	Bloomington # 41212
UAS-miR-980	w [1118]; + ; UAS-miR-980	Bloomington # 41191
UAS-miR-982	w [1118]; + ; UAS-miR-982	Bloomington # 41192
UAS-miR-983-1	w [1118]; + ; UAS-miR-983-1	Bloomington # 41194
UAS-miR-984, 983-1, 983-2	w [1118]; + ; UAS-miR-984, 983-1, 983-2	Bloomington # 41217
UAS-miR-984	w [1118]; + ; UAS-miR-984	Bloomington # 41224
UAS-miR-985	w [1118]; + ; UAS-miR-985	Bloomington # 41213
UAS-miR-986	w [1118]; + ; UAS-miR-986	Bloomington # 41218
UAS-miR-987	w [1118]; + ; UAS-miR-987	Bloomington # 41195
UAS-miR-988	w [1118]; + ; UAS-miR-988	Bloomington # 41196
UAS-miR-989	w [1118]; + ; UAS-miR-989	Bloomington # 41219
UAS-miR-992	w [1118]; + ; UAS-miR-992	Bloomington # 41130
UAS-miR-993	w [1118]; + ; UAS-miR-993	Bloomington # 41197
UAS-miR-994	w [1118]; + ; UAS-miR-994	Bloomington # 41198
UAS-miR-995	w [1118]; + ; UAS-miR-995	Bloomington # 41199
UAS-miR-999	w [1118]; + ; UAS-miR-999	Bloomington # 41200
UAS-miR-9a	w [1118]; + ; UAS-miR-9a	Bloomington # 41138
UAS-miR-9b	w [1118]; + ; UAS-miR-9b	Bloomington # 41131
UAS-miR-9c	w [1118]; + ; UAS-miR-9c	Bloomington # 41139
UAS-miR-iab-4	w [1118]; + ; UAS-miR-iab-4	Bloomington # 41173
UAS-miR-5,4,286	; JB 43B/CyO UAS-DSRed-mir-5/4/286; +	Lai lab (Lai et al. 2005)
UAS-bantam	; + ; bantam UAS-C/TM6	Cohen lab (Thompson & Cohen 2006)
UAS-miR-6-1, 6-2, 6-3	;UAS-mir-6-1,2,3 (3A)/SM; +	Lai lab (Lai et al. 2005)
UAS-miR-7	w[*]; + ; UAS-mir7[140]	Alonso lab
Ubx>GAL4[LDN]	y w; + ; Ubx>GAL4[LDN]/TM6b	Sa´nchez-Herrero lab (Navas et al. 2006)

UAS-mCherry	yw/w; UAS-mCherryNLS/ (CyO); MKRS/TM6B	Affolter lab (Caussinus et al. 2008)
sd>GAL4	w* P{GawB}sdSG29.1	Bloomington # 8609
UAS-mCherry.K10	yw; UAS-mCherry.K10 (Line 2); +	Alonso lab (Thomsen et al. 2010)
UAS-mCherry.long	yw; UAS-mCherry.long (Line 2): +	Alonso lab (Thomsen et al. 2010)
Oregon Red	+; +; +;	Alonso lab
w1118	w [1118]; + ; +	Alonso lab
elav>Gal4	w[*]; + ; elav>Gal4	Bloomington # 8760
Δmir-92a,b (deletion stock)	w[1118]; Df(3R)BSC321/TM6C, Sb[1] cu[1]	Bloomington # 24909
Δmir-252	y^{*} w^{*}; P{$w^{+mW.hs}$=FRT(w^{hs})}2A P{$ry^{+t7.2}$=neoFRT}82B PBac{SAstopDsRed}LL04028 P{$y^{+t7.7}$ $ry^{+t7.2}$=Car20y}96E / TM6B, Tb^{1}	DGRC # 140830
miR-310c>GAL4	y^{*} w^{*}; P{$w^{+mW.hs}$=GawB}NP5941 / CyO, P{w^{-}=UAS-lacZ.UW14}UW14	DGRC # 113798

2.2 Haltere cuticle preparation, mounting and imaging

Haltere cuticle preparation and dissection were performed as described in the study of Navas et al. (Navas et al. 2006) with a subtle modification. Briefly, newly enclosed flies, 1-2 days after enclosing, were rolled in 1:3 glycerol/ethanol solution for 1 hour and macerated in 10% KOH at 65°C for about 1 hour. Samples were then washed and kept in 1:3 glycerol/ethanol solutions. Haltere cuticle dissection and mounting were carried out in 70% glycerol/PBTriton [0.3% Triton X-100 in 1XPBS (Phosphate buffered saline)]. All samples were imaged using Leica DMRB microscope.

2.3 Haltere appendages preparation and scanning electron microscope (SEM)

Newly enclosed adult flies (1-2 days after enclosing) were kept in an empty vial for 1 hour to clean the tissues. A series of 70%, 90%, 100% and 100% ethanol solutions were used to dehydrate flies subsequently for 10 minutes each. Flies were washed twice (1 hour each) with Hexamethyldisilazane (Sigma Aldrich) as an alternative to critical point drying. Samples were evaporated in the hood overnight for the future mounting. The treated flies were then mounted on the

carbon tape and coated with the gold. All samples were imaged with Leo stereo scan 420 scanning electron microscope.

2.4 Haltere imaginal discs dissection and fixation

Later 3rd instar larvae were picked and kept in 1XPBS on ice for 10 minutes to clean the guts. Samples were dissected in ice-cold 1XPBS by using forceps to tear the larvae apart from the middle and turn the anterior part of larvae inside out (Sullivan et al. 2000). Therefore, haltere imaginal discs attached to the body wall of the larvae were exposed to the solution. Samples were fixed in 4% formaldehyde in 1XPBS (diluted from 10% formaldehyde, Polysciences, Inc.) at room temperature for 20 minutes and stored in 100% methanol at -20°C.

2.5 Embryo Collection and Fixation

Embryos were collected at 25°C on apple juice agar plates with a sprinkling of yeast. Samples were dechorionated in 50% bleach for 5 minutes and fixed for 22 minutes in 1:1 heptane / (259µl 10% ultrapure Formaldehyde and 241µl 1XPBS) at room temperature. The vitelline membrane was removed by vigorously shaking the embryos in 1:1 heptane/methanol for at least 2 minutes. To remove the formaldehyde, samples were then rinsed with 100% methanol several times. Embryos were stored in 100% methanol at -20°C for later use.

The above procedure was used for the embryo preparation of immunohistochemistry analysis. In the case of fluorescent *in situ* hybridisation, the fixation procedure is the same, but was carried out in RNase-free condition.

2.6 Immunohistochemistry

Immunohistochemistry was operated following standard procedures (Rogulja-Ortmann *et al*. 2014). Samples were incubated with primary antibody at 4°C overnight on the shaker. Following extensive washing, samples were incubated

with secondary antibody at room temperature for two hours. For the antibody staining of embryos, 10 minutes of 0.001% sodium borohydride (ACROS) in 0.1% PBTween treatment was added to reduce the auto-fluorescence. The antibodies used in this thesis are listed as below (Table 2.2, 2.3). Leica DM6000 fluorescent microscope, Leica TCS SPII confocal microscope, Zeiss Axiophot confocal microscope and Leica SP8 confocal microscope were used for the fluorescent samples imaging. All images were analysed by ImageJ.

Table 2.2 Primary antibodies used in this study.

Name	Host	Concentration	Origin
anti-Ubx	mouse	1:20	Gift from Rob White, University of Cambridge
anti-DsRed	rabbit	1:500	clontech
anti-repo	rabbit	1:500	a gift from Gerd Technau, University of Mainz
anti-engrailed	rabbit	1:100	Santa Cruz Biotechnology
anti-GFP	rabbit	1:300	Molecular probes
anti-repo	mouse	1:20	DSHB

Table 2.3 Secondary antibody used in this study.

Name	Concentration	Origin
anti-mouse Alexa 488	1:500	Life technologies
Dapi	1:500	Life technologies
anti-mouse Alexa 555	1:500	Life technologies
anti-rabbit Alexa 488	1:500	Life technologies
anti-rabbit Alexa 555	1:500	Life technologies
anti-rabbit Alexa 647	1:100	Life technologies
anti-mouse Alexa 647	1:400	abcam

2.7 Embryos collection and RNA extraction

Flies were kept at 25°C in large collection cages with apple juice plates supplemented with a small amount of yeast to trigger flies to lay eggs. Embryo collection was carried out by two-hour time windows from 0 to 24 hours with one-hour prelay. Samples were dechorionated in 50% bleach with DEPC

treated RNase-free water for 5 minutes and rinsed with DEPC treated RNase-free 1XPBS to wash out the bleach. The dechorionated embryos were moved to 1.5ml RNase-free Eppendorf tubes and homogenized in 50µl TRIzol® Reagent. After fully homogenization, samples were added a final volume of 500µl Trizol for extraction. Total RNA extraction was performed using TRI Reagent solution (Ambion) according to the manufacturer's instruction. The RNA pellet was suspended in 50µl DEPC treated water and stored at -80°C until use. RNA concentration was quantified by Nanodrop 2000 spectrophotometer (Thermo Scientific).

2.8 Reverse transcription polymerase chain reaction (RT-PCR)

2ug of total RNA was treated with DNase I (New England Biolabs) at 37°C for 10 minutes to eliminate genomic DNA. Treated RNA samples were used for cDNA synthesis with RETROscript Kit (Ambion) according to the manufacturer's protocol. cDNA was stored in -20°C until use.

1µl cDNA was used for 25µl final volume PCR reaction according to standard protocol. 29 cycles were used for amplification with the annealing temperature of 58°C. Primers used were designed by Primer3 and listed in Table 2.4. The ethidium bromide gel quantification was done by ImageJ.

Table 2.4 Primers used for semi-quantitative RT-PCR in this thesis.

Gene	Forward primer (5' to 3')	Reverse primer (5' to 3')	Amplicon length	Source
***Ubx* short 3' UTRs**	GAAATGACGCGGAGACAGAT	AATCTGCGCTCCTTCCACTA	236	(Thomsen et al. 2010)
***Ubx* long 3' UTRs**	GAACGAAGGCAGATGCAAAT	GGTAAGTGGTCGGATGCAGT	225	(Thomsen et al. 2010)
Rp49	CCAGTCGGATCGATATGCTAA	TCTGCATGAGCAGGACCTC	268	(Thomsen et al. 2010)
miR-92a	CGATGCTCCTATTGTTCGCC	ACAAGAAGCAAACTCACCTGT	377	This study
miR-92b	ATCCGAGATGTGAGTGCAGT	TGGGCGTCATTTTGGAGCT	297	This study
miR-252	ATCATTGCAGAACATGGGGC	CACTGAGCAGCGTGTATTCT	222	This study

2.9 TaqMan RT-PCR

This work is a collaboration project with our previous lab member Dr. Richard Kaschula. RNA samples were extracted from wild-type haltere imaginal discs at the white pre-pupa stage, which were dissected by Dr. Richard Kaschula and Dr. Ana Bomtorin, a visiting student from Universidade de Sao Paulo, Ribeirao Preto, Brazil. Richard extracted total RNA from about 800 haltere imaginal discs collected from Or-R white pre-pupa using Tri Reagent (Ambion) following the manufacturer's protocol. 2µl of glycogen (Ambion) was added to each sample during -80°C RNA isopropanol precipitation to increase yield.

I carried out the TaqMan RT-PCR procedure. 5µl of 2ng/ µl total RNA was used for TaqMan RT-PCR by TaqMan miRNA assays (Life technologies) according to manufacturer's standard protocol. U27 is an artificial assay used as a control. The reverse transcription was operated on an iCycler BioRad PCR machine and the PCR reaction was performed in LightCycler 480 II (Roche). miRNA assays used in this thesis were listed as the followings (Table 2.5).

Table 2.5 miRNA assays used in this thesis for TaqMan RT-PCR

Assay ID	Assay Name	Assay Mix Concentration	Reporter 1 Dye	Reporter 1 Quencher	Context Sequence
1752	U27	20x	FAM	NFQ	GTTCTGTGATGTCAAACCAATAGAC AAGCATATAACCGAACAATCATGTT GATTTTCACACGACTGAGC
313	*dme-miR-304*	20x	FAM	NFQ	UAAUCUCAAUUUGUAAAUGUGAG
006442_mat	*aae-miR-137*	20x	FAM	NFQ	UAUUGCUUGAGAAUACACGUAG
008279_mat	*dme-miR-980-3p*	20x	FAM	NFQ	UAGCUGCCUUGUGAAGGGCUUA
004649_mat	*dme-miR-190-5p*	20x	FAM	NFQ	AGAUAUGUUUGAUAUUCUUGGUUG
277	*dme-miR-13b*	20x	FAM	NFQ	UAUCACAGCCAUUUUGACGAGU
323	*dme-miR-313*	20x	FAM	NFQ	UAUUGCACUUUUCACAGCCCGA
332	*dme-let-7*	20x	FAM	NFQ	UGAGGUAGUAGGUUGUAUAGU
583	*hsa-miR-9*	20x	FAM	NFQ	UCUUUGGUUAUCUAGCUGUAUGA
464458_mat	*bmo-miR-252*	20x	FAM	NFQ	CUAAGUACUAGUGCCGCAGGAG
005481_mat	*dme-miR-1006-3p*	20x	FAM	NFQ	UAAAUUCGAUUUCUUAUUCAUAG
245576_mat	*dme-miR-iab-4-5p*	20x	FAM	NFQ	ACGUAUACUGAAUGUAUCCUGA
462046_mat	*dme-miR-1012-3p*	20x	FAM	NFQ	UUAGUCAAAGAUUUUCCCCAUAG
465034_mat	*dme-miR-375-3p*	20x	FAM	NFQ	UUUGUUCGUUUGGCUUAAGUUA
476768_mat	*dme-miR-282-5p*	20x	FAM	NFQ	UAGCCUCUACUAGGCUUUGUCUGU

2.10 RNA probe synthesis

Genomic DNA extraction was performed from adult Or-R flies by using the DNA extraction kit (Transgene) according to the manufacturer's instructions. Primers for templates of RNA probes were designed by Primer3 and list in Table 2.6. PCR reaction products were cloned into pGEM-T-easy (Promega) using manufacturer's procedure. Plasmids were linearized and purified by QIAquick PCR purification kit (QIAGEN) according manufacturer's protocol. The concentration of linearized plasmid was measured by Nanodrop 2000 spectrophotometer (Thermo Scientific).

Table 2.6 Primers used to synthesize the FISH probes in this thesis

gene	Forward primer (5' to 3')	Reverse primer (5' to 3')	Probe Length	Source
***Ubx* long 3' UTRs**	CGTGTGTGTGTCCCGATAAT	TCCACATTCTCACTGGTTGC	819	(Thomsen et al. 2010)
***Ubx* intron 3A**	AAGGGTACGACCACTGCAAC	GCGGTACCTCGGACAATTTA	843	(Rogulja-Ortmann et al. 2014)
***Ubx* intron 3B**	AGCCGGCATCCAGACTACTA	GCATACCAGAGACCCAGCAT	826	(Rogulja-Ortmann et al. 2014)
***Ubx* intron 3C**	ATTGGCTACCCATCTGCAAC	TGCTACCCCTCTTCCTACCA	844	(Rogulja-Ortmann et al. 2014)
miR-92a	GTTGGCAAATCCTGTCCAGT	TGTGGAATTGTGTGGCTGAT	819	This study
miR-92a	CTATTTGCGTTTCCGCTCTC	CCATCAACAGGTAGGCAGGT	937	This study
miR-92a	CCCCAAAACAGGTTCCTTCG	TGATGGCTATTTCTCCGCCT	872	This study
miR-252	TGGAGGCATCTTTGAGCTGT	TACAGTGGCGTTTCTTCCCT	536	This study
miR-92b	TAATTCGAGCGGAAACTGCT	AAGATGATGGATGCCCTCTG	711	This study
miR-92b	TTCTCCATATGGCCGGATTA	TCGTCACTTCTACGCCTTCA	884	This study
miR-92b	TTCTCCATATGGCCGGATTA	AAGATGATGGATGCCCTCTG	909	This study

According to the orientation of the PCR amplicon in the plasmid, 1μg purified templates were used to synthesize RNA probes by either T7 or SP6 RNA polymerase (Roche) using DIG RNA labeling mix (Roche) corresponding to manufacturer's practice with slight modification. DNA templates were removed by DNase I (New England Biolabs) and RNA probes were precipitated with 2.5μl 4M lithium chloride and 75μl pre-chilled 70% ethanol solution at -80°C for 2 hours or overnight. RNA probes were centrifuged at 4°C for half an hour and air-dried at room temperature. RNA pellets were re-suspended in 50μl hybridisation buffer [50% formamide (ACROS), 5x SSC (Sigma), 100ug/ml salmon sperm DNA (Invitrogen), 0.1% PBTween (1XPBX RNase-free, 0.1% Tween-20, Sigma)] and stored in -20°C.

2.11 Fluorescent *in situ* hybridisation (FISH)

FISH for *miR-92a*, *miR-252*, *miR-92b* and *miR-10* were performed as described (Rogulja-Ortmann *et al*. 2014) with some modifications. Fixed embryos were treated with 3% H_2O_2 (30% H_2O_2, ACROS) in methanol to quench endogenous HRP, and then rehydrated in 0.1% PBTween with 20 minutes wash for three times. Samples were incubated for 10 minutes with 0.001% sodium borohydride (ACROS) in 0.1% PBTween to reduce autofluorescence. Embryos were pre-hybridised in hybridisation buffer for 1 hour. RNA probes were denatured in heat blocker for 5 minutes and incubated with pre-hybridised embryos at 55°C overnight. All steps were carried out in RNase-free condition.

For the DIG labelled RNA detection, embryos were blocked in TNB solution [(0.1m Tris pH 7.5 (Fisher), 0.15M NaCl (Fisher), 0.5% blocking reagent (Roche)) for 30 minutes and incubated with 1:500 anti-DIG-POD diluted in amplification solution (Roche) for 2 hours at room temperature. The fluorescence signal was detected by either FITC TSA plus amplification kit (Perkin Elmer) or Cy3 TSA plus amplification kit (Perkin Elmer) according to manufacturer's instructions.

For double FISH, two RNA probes were labelled with either DIG or FITC independently and detected using coordinated antibody and amplification kit: DIG labelled RNA was detected by anti-DIG-POD with Cy3 TSA plus amplification kit (Perkin Elmer), while FITC labelled RNA was recognized by anti-FITC-POD and amplified by FITC TSA plus amplification kit (Perkin Elmer). 20 minutes treatment with 0.3% H_2O_2 in methanol was added between two amplifications to avoid cross detection.

2.12 Bioinformatic programs and predictions

Here, I showed the features of ten main used *Drosophila* miRNA prediction programs and the prediction methods used in this study for each program. Unless the prediction program provided the miRNA and 3' UTR sequence,

Drosophila melanogaster miRNA sequences for prediction were acquired from miRbase 21 and the longest version of *Ubx* 3' UTR was download from Flybase as the miRNA targeted sequence.

2.12.1 Programs with local prediction

miRanda (Enright et al. 2003), RNA22 (Miranda et al. 2006) and RNAhybrid (Krüger & Rehmsmeier 2006; Rehmsmeier *et al*. 2004) provide the executable resources for local prediction. The detailed features of each programs are described as followings.

2.12.1.1 miRanda

The algorithm of miRNA target predictions for miRanda (Betel et al. 2008; Enright et al. 2003) was established by considering both sequence complementarity using a position-weighted local alignment algorithm and free energies of RNA-RNA interaction. This algorithm also considers the conservation of target sites in related genomes including *Drosophila melanogaster, Drosophila pseudoobscura and Anopheles gambiae*. Additionally, an optional feature, called mirSVR, which is a new machine learning method for ranking miRNA target sites by a down-regulation score, could be applied to identify experimentally determined non-canonical and non-conserved sites. This feature dramatically increased the number of predicted miRNA target sites (Betel et al. 2010). In addition, this program takes into account the cooperative regulation of miRNAs and miRNA target multiplicity.

2.12.1.2 RNA22 version 1.0

RNA 22 (Miranda et al. 2006) uses the *Teiresias* algorithm to identify the pattern of mature miRNA sequences. Then it uses a second-order Markov chain to estimate the statistical significance of each pattern and discards patterns with estimated log-probability above -38. The reverse complements of identified patterns are generated and matched to the sequence of the putative mRNA target. The selected sequences that surpass the statistical threshold are defined

as target islands, and are then recognized by miRNAs depending on complementary paring and free energy calculation. Identification of the target sites even without knowing the identity of the targeting miRNA allows us to capture target sites for miRNAs that are not currently recognized. By using a luciferase-reporter experimental approach to estimate the utility of this program, RNA22 not only provides a good performance of its miRNA target algorithm, but also strongly suggests that some miRNAs may have thousands of target mRNAs.

2.12.1.3 RNAhybrid

RNAhybrid (Rehmsmeier et al. 2004) is a miRNA prediction tool based on the most energetically favorable hybridisation sites of a short RNA in a longer mRNA sequence, so that the short RNA sequence is matched to the best fitting region of the long RNA. The design of this program avoids self-pairing between target nucleotides or among miRNA nucleotides. Additionally, this program calculates P-values depending on extreme value distributions of length-normalized energies. Additional tools that aid in this software's function have been developed, such as RNAcalibrate and RNAeffective. RNAcalibrate is a tool used for calibrating the performance of minimum free energy hybridisation by RNAhybrid, while RNAeffective is used to determine the effective number of orthologous miRNA targets. A new update for this program also added the function of the possibility to disallow G:U base pairs in the seed region and a seed-match speed-up (Krüger & Rehmsmeier 2006). This program is suitable to perform genome-wide predictions of non-canonical target sites and could be applied to do miRNA target prediction in both 3' UTR and CDS of mRNA.

2.12.1.4 Bioinformatic predictions for miRanda, RNA22 and RNAhybrid

The miRNA predictions for miRanda, RNA22 and RNAhybrid use local predictions, which are allowed to input the particular miRNAs and mRNAs of interest. All 182 miRNAs of self-input were downloaded from miRbase release 21 and the *Ubx* 3' UTR sequence was retrieved from flybase. miRNA target prediction for miRanda, I run the miRanda-3.3a. For RNA22, I used

remoteRNA22v2. For RNAhybrid, I downloaded the RNAhybrid Windows (32 Bit) binary package from the website (Table 1.1) and performed miRNA prediction using the code for accurate P-value calculation of length normalized minimum free energies.

2.12.2 PITA

PITA (Kertesz et al. 2007) considers the target site accessibility by calculating miRNA and mRNA interactions. In addition, this program takes the miRNA seed-3' UTR pairing into account, and calculates the difference between the energy gained from miRNA binding to mRNAs (duplex) and the energy lost when the mRNA target-site nucleotides are locally unpaired, facilitating miRNA access. The minimum length for the seed paring is 6nt, and one G (guanine):U (uracil) wobble pair is tolerable when having 8nt-long sequence paring in the seed sequence. This miRNA targeting prediction program also considers the upstream and downstream flanking sequences of target sites.

The miRNA prediction of PITA was done using the web tool default setting for *Drosophila* (Table 1.1). All 182 miRNAs of self-input were downloaded from miRbase release 21 and the 3' UTR sequence of *Ubx* was retrieved from flybase. The minimum seed size is 8 nt, and allowed maximally a single G: U and a single mismatch in the seed region. I also used 3 upstream and 15 downstream flank setting and target sites with the delta delta G below -0 were selected (Kertesz et al. 2007).

2.12.3 ComiR

ComiR (Coronnello et al. 2012) uses a combination of binding models, incorporating the improved predictions of four target programs into a single probabilistic score using ensemble learning. Each algorithm is run separately, predicting all binding sites for each miRNA on the 3' UTR of a given mRNA. A novel feature of ComiR is that this miRNA target-prediction tool considers the expression level of each miRNA before determining the regulatory potential of combined miRNAs to one particular mRNA. After incorporating miRNA

expression data, this program then combines the individual target scores produced by PITA and miRanda using a Fermi-Dirac equation (FD score) and combines the individual target scores of TargetScan and mirSVR by weight sum (WSUM score), respectively. A support vector machine (SVM) with linear kernel trained on the *Drosophila* RISC IP dataset was used to combine the final scores of the four aforementioned miRNA targeting-prediction tools for each target mRNA (Coronnello & Benos 2013).

The miRNA prediction of ComiR was carried out with default setting on the program website (Table 1.1) without considering the miRNA expression level (Coronnello & Benos 2013). The threshold of miRNA-mRNA interaction was set with comiR score>=0.9027, above that of *miRNA-iab-4* and *miRNA-iab-8*, since these two miRNAs had been validated experimentally (Tyler et al. 2008; Ronshaugen et al. 2005). All self-input miRNAs were downloaded from miRbase release 20 and the 3' UTR sequence of *Ubx* was obtained from flybase.

2.12.4 DIANA-microT web server v5.0

The DIANA-microT algorithm for miRNA target prediction is specifically performed on both a positive and a negative set of miRNA Recognition Elements (MREs) that is defined by Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP) data. An independent prediction model based on MREs was built for CDSs and 3' UTRs, and combined for the final calculation of miRNA and target mRNA interaction score. High-throughput proteomics data is used for the performance test. By using this prediction algorithm, many protein-coding genes were found to have potential targets of miRNAs exclusively in the CDSs, although this miRNA target program provides a model to predict miRNA targeting in both CDSs and 3' UTRs (Reczko et al. 2012). DIANA-microT-CDS incorporates miRBase version 18 and Ensembl version 69, and is also suitable for the miRNA target prediction in *Homo sapiens, Mus musculus, Drosophila melanogaster and C.elegans* (Paraskevopoulou et al. 2013).

For DIANA, I used DIANA-microT web server v5.0 (Table 1.1) to do the miRNA prediction and the threshold was miTG score>=0.7 according to the default setting of the website. The miRNA input used was obtained from miRBase build 13 and the mRNA input was downloaded from Ensembl build 54 (Reczko et al. 2012).

2.12.5 EMBL

This is the only miRNA target prediction program designed for *Drosophila melanogaster* exclusively among all ten miRNA prediction programs. The prediction of EMBL is based on pairing rules from an *in vivo* systematic experimental study. The algorithm considers both 5' dominant sites and 3' compensatory sites and only the evolutionarily conserved target sites are taken into account according to genome comparison across insect species. A highly sensitive and specific performance of this program was suggested by the authors, based on experimental data (Brennecke et al. 2005). Files can be downloaded for either individual miRNA target predictions or miRNA target predictions for all available miRNAs, on the website with full annotations such as CG-ID, gene name, Flybase-ID, GO-terms, validated (Real) or predicted (PRED) UTR, score of best site, total score of all sites and number of sites.

miRNA prediction by EMBL, I used the version of prediction updated in 2005 by downloading all predictions results for *Ubx* from their website (Table 1.1) with the default setting. This prediction tool contains 55 mature miRNAs from Rfam and the *Ubx* 3' UTR was downloaded from the Berkeley Drosophila Genome Project (BDGP) (Brennecke et al. 2005).

2.12.6 miRecord

miRecord (Xiao et al. 2009) includes two kinds of miRNA-mRNA interactions data: Validated targets and Predicted targets. The "Validated targets" consist of a large database of experimentally validated miRNA and mRNA interactions. 2705 records of miRNA-mRNA interactions between 644 miRNAs and 1901 target genes in 9 animal species has been reported as of the update on 27 April

2013 on the website of this program (Table 1.1). The “Predicted targets” part of miRecord integrates the prediction results from 11 miRNA prediction programs, including DIANA-microT, miRanda, PicTar, PITA, RNA22, TargetScan, MicroInspector, MirTarget2, miTarget, NBmiRTar, and RNAhybrid. Each miRNA target prediction program runs separately and the final prediction results of each program are summarized in a list, which greatly improves the efficiency of miRNA target prediction.

miRNA prediction of miRecord used the web tool default settings (Table 1.1). In total, 152 mature miRNAs were used for the target prediction. I counted miRNAs target interactions predicted by at least one miRNA prediction program (Xiao et al. 2009).

2.12.7 PicTar

PicTar (Krek et al. 2005) uses multiple alignments for 3’ UTR nucleotide (nt) sequences and information on co-expressed mature miRNA to nuclMap either by perfect nucletide match (7nt starting at position 1 or 2 of the 5’ end of the miRNA) or imperfect nucletide match (containing no more than one nucleotide bulge, G:U wobble or mismatch) in the aligned 3’ UTR sequences. The binding stability of putative miRNA and mRNA interaction is calculated as free energy to filter out miRNA-3’ UTR interactions that are not thermodynamically possible. Then, PicTar uses Hidden Markov Model (HMM) to calculate the maximum likelihood score (PicTar score) and individual score calculated for each single 3’ UTR alignment from a set of species were combined to obtain the final score. PicTar is useful for miRNA target predictions of both individual as well as combinatorial miRNA binding.

PicTar miRNA target prediction was carried out with the default settings (Table 1.1) with the filter of single miRNA target predictions (setting S1: high sensitivity). All 70 mature miRNAs input were obtained from Rfam and the 3’ UTR sequence was retrieved from the UCSC Genome browser database (Grün et al. 2005).

2.12.8 TargetScanFly Release 6.2

TargetScan (Kheradpour et al. 2007; Lewis et al. 2003; Lewis et al. 2005) predicts miRNA targets depending on the conserved 8-mer (8 nucleotides of the mRNA pair with the 'seed' in the 5′ end of the miRNA) and 7-mer sites that are perfectly matched to the seed region of the miRNAs. Other features such as an A (adenine) in position 1 of the target site, sequence complementarity in position 8, AU flanking and the distance between the target site and the end of 3’ UTR, are also taken into account. The conservation cutoffs are measured by branch length score (BLS), which is calculated by evaluating the total evolutionary branch length of the phylogenetic tree over which the motif appears conserved. However, poorly conserved miRNA target sites are also provided as a complement.

The miRNA prediction by TargetScan was also carried out with default setting (Table 1.1) based on TargetScanFly Release 6.2. In total, 148 miRNAs were used for miRNA prediction (Ruby et al. 2007) and the 3’ UTR of the longest annotated *Ubx* from the UCSC Genome browser database (https://genome.ucsc.edu/) was used for the prediction.

2.13 Statistical analysis

Statistical analyses were carried out in Prism GraphPad 6.0 software package. I either use unpaired, non-parametric with Mann-Whitney test or unpaired, parametric with Welch’s correction, depending on the feature of the data. (P-values’ possibility: $P>0.05$ (non-significant; n.s.), $P<0.05$ (*), $P<0.01$ (**), $P<0.001$ (***), $P<0.0001$ (****)).

CHAPTER 3 Identification of *Hox* regulatory miRNAs via a genetic screen in the developing haltere

3.1 Introduction

Increasing evidence suggests that miRNA-mediated post-transcriptional regulation is more relevant to gene expression control than previously thought (Bartel 2004; Eulalio et al. 2008; Filipowicz et al. 2008; Kloosterman & Plasterk 2006; Nilsen 2007). Bioinformatics and experimental data suggests that miRNAs regulate more than 30% of human protein-coding genes (Nilsen 2007; Filipowicz et al. 2008) and previous studies have shown that miRNAs regulate gene expression in different species, which indicates that miRNAs are prevalent and conserved regulators in the development of multicellular organisms (Cordes et al. 2009; Hwang et al. 2014; Tan et al. 2013; Reinhart et al. 2000; Arif et al. 2013). A salient study from our lab has shown that different *Ubx* 3' UTR isoforms are differentially expressed during *Drosophila* embryonic CNS development (Thomsen et al. 2010). The alternative processing of *Ubx* 3' UTR leads to *Ubx* mRNA isoforms that contain different sets of miRNA targets suggesting a prevalent role for miRNA regulation during neuronal development (Bartel & Chen 2004). Despite the importance of miRNA regulation in modern Biology and Biomedicine, fundamental questions regarding their function remain open.

miRNA target prediction using bioinformatics programs appears as a reasonable first approach to tackle this question. Since 2003, several groups have used bioinformatics approaches generating algorithms to predict miRNA targets (Stark et al. 2003; Enright et al. 2003; Rajewsky 2006; Kiriakidou et al. 2004). However, animal miRNAs find their targets by partial complementarity (Yekta et al. 2004; Davis et al. 2005), making target prediction very difficult. To complement target prediction algorithms, experimental identification of miRNA targets through ectopic expression of individual miRNAs emerges as a promising approach to corroborate miRNA-target interactions (Bejarano et al. 2012; Schertel et al. 2012; Szuplewski et al. 2012).

The *Hox* gene *Ubx* encodes a homeodomain transcription factor that controls the genetic networks underlying haltere development in *Drosophila* (Weatherbee et al. 1998). Among the known 256 *Drosophila* miRNAs (miRBase 21) (Kozomara & Griffiths-Jones 2014), only *miR-iab-4* and *miR-iab-8* have been proven to regulate *Ubx* expression during *Drosophila* development (Bender 2008; Ronshaugen et al. 2005; Stark et al. 2008; Thomsen et al. 2010; Tyler et al. 2008). These two miRNAs are encoded within the *Hox* cluster (Tyler et al. 2008; Stark et al. 2008; Ronshaugen et al. 2005). Given that current estimates support the existence of 256 miRNAs precursors which can be processed to produce 466 mature miRNAs in *Drosophila melanogaster* (miRBase 21) (Kozomara & Griffiths-Jones 2014), it seems highly plausible that miRNAs encoded outside the *Hox* cluster might also be able to regulate *Hox* gene expression. Therefore, I wonder whether any of the remaining 254 miRNAs present in *Drosophila* play any role in the regulation of *Ubx*. Are there any other *Ubx* regulatory miRNAs in *Drosophila*? If so, how could there other miRNAs be identified?

Different *Ubx* mutant alleles cause different expression level of *Ubx* in the haltere imaginal discs (Figure 1.4B), which is correlated with their varying severity of the homeotic transformation in the adult appendage (Bender et al. 1983). Given that a reduction in *Ubx* expression levels causes well-established morphological changes in the adult haltere, in this chapter, I present an experimental approach to detect miRNAs able to regulate the *Hox* gene *Ubx*, one of the *Hox* genes whose post-transcriptional RNA regulation is currently best understood (Mallo & Alonso 2013; Rogulja-Ortmann et al. 2014; Reed et al. 2010; Thomsen et al. 2010). I developed a genetic screen by using ectopic expression of miRNAs in haltere imaginal discs based on the UAS-GAL4 system (Duffy 2002; Brand & Perrimon 1993) and score adult haltere phenotypes to identify all potential miRNAs that may repress *Ubx* expression.

Across all the 96 miRNAs and 10 miRNA clusters tested, I classified the miRNAs into four phenotypic classes, and found that around 30% miRNAs potentially regulate *Ubx* expression. In order to understand the molecular basis underlying the haltere phenotypes generated and miRNA expression level in

pre-pupal haltere imaginal discs, I used a combination of RNA Sequencing (RNA-seq) data from a previous lab member, Dr Richard Kaschula (The Francis Crick Institute, London, UK) together with my own TaqMan RT-PCR experiments to validate the RNA-seq approach, and to compare expression levels of miRNAs in different phenotypic classes. Statistical analysis revealed that there is no correlation between the strength of haltere phenotypes and miRNA expression level in pre-pupal haltere imaginal discs. So haltere morphological changes induced by ectopic expression of miRNAs appears to be the consequences of the biological function of the miRNAs, instead of unspecific targeting due to high miRNA expression.

3.2 Results

3.2.1 Using the *Hox* gene *Ubx* to identify regulatory miRNAs

Ectopic expression of individual miRNAs can induce specific phenotypes that recapitulate target gene inactivation, providing a morphological readout that can guide our functional study. Several groups have carried out miRNA gain-of-function screen to study miRNA-mRNA regulation showing that it is a useful tool to complement loss-of-function approaches (Bejarano et al. 2012; Schertel et al. 2012; Szuplewski et al. 2012). It has been proven that miRNA gain-of-function phenotypes relate to miRNA loss-of-function phenotypes. For example, the ectopic expression of *dme-miR-iab-4* and *dme-miR-iab-8* leads to a decrease in Ubx expression levels in the haltere imaginal discs, which in turn causes haltere-to-wing morphological transformations in the adult fly (Ronshaugen et al. 2005; Stark et al. 2008). This is in agreement with the increase of Ubx protein levels when lacking the expression of *dme-miR-iab-4* and *dme-miR-iab-8* in *Drosophila* CNS (Bender 2008; Thomsen et al. 2010). Thus, miRNA ectopic expression emerges as a valid and promising approach to find potential miRNA-mRNA interactions.

The *Drosophila* haltere provides a very sensitive readout system to detect fluctuation in the expression level of *Ubx*. Following the foundational

observations made by Calvin Bridges and Thomas Hunt Morgan in 1923 reporting the isolation of the first *bithorax* mutation (Bridges & Morgan 1923), in 1978, Edward B. Lewis described a series of BX-C mutants which caused a haltere-to-wing transformation in line with a homeotic change in appendage specification (Lewis 1978). Reduction of *Ubx* expression in the haltere leads to bigger halteres and the formation of ectopic bristles on the anterior part of the haltere (Stark et al. 2008). This feature, i.e. the morphological modification of the haltere caused by a reduction in *Ubx* expression provides an excellent system to observe how ectopic expression of miRNAs alters the *Ubx* expression level (Figure 3.1A). Thus, using the UAS-GAL4 system to study gain-of-function for miRNAs is an efficient experimental approach to detect other miRNAs able to regulate the *Hox* gene *Ubx*.

In this work, I used 106 UAS-miRNA transgenes lines, including 96 miRNAs and 10 miRNA clusters (Table 2.1) that were available for public use to carry out the genetic screen. I used the *Ubx*-Gal4^{LDN} driver for all miRNA ectopic expressions (Figure 3.1B). miRNA overexpressed by *Ubx*-Gal4^{LDN} can lead to exogenous miRNA expression in the haltere pouch, larval CNS and in the patches of cells in the third leg disc (Navas et al. 2006). It is important to note that the *Ubx*-Gal4^{LDN} driver is a GAL4 insertion within *Ubx* regulatory regions leading to a mild *Ubx* mutant phenotype (Navas et al. 2006). Thus, the genetic screen is an enhancer screen: upon ectopic expression of miRNAs that repress *Ubx* expression in an *Ubx*-Gal4^{LDN} background, I expect an enhancement of effects on haltere morphology. Therefore, the ‘sensitised’ genetic background provided by the *Ubx*-Gal4^{LDN} background should increase our capacity to detect miRNA effects on *Ubx* expression in the haltere.

Figure 3.1 illustrates the strategy of the genetic screen. The *Ubx*-Gal4^{LDN} fly expresses the yeast transcription activator protein GAL4 in the haltere pouch. This protein then binds to the upstream activation sequence (UAS) that acts as an enhancer in the miRNA transgene lines to activate exogenous miRNAs expression in the haltere pouch. Depending on the transgenic flies utilized, there will be different kinds of mature miRNAs expressed in the haltere pouch. For example, the transgenic fly expressing a single miRNA can produce

different combinations of functional mature miRNAs, including 3' arm, 5' arm or both 3' and 5' arm depending on the maturation of the pre-miRNA. For the transgenic fly of a miRNA cluster, I can have all functional mature miRNAs produced from the pre-miRNAs in this cluster. The *Drosophila* adult haltere is then used as readout to measure to what extent the *Ubx* expression level is changed by ectopic expression of miRNA.

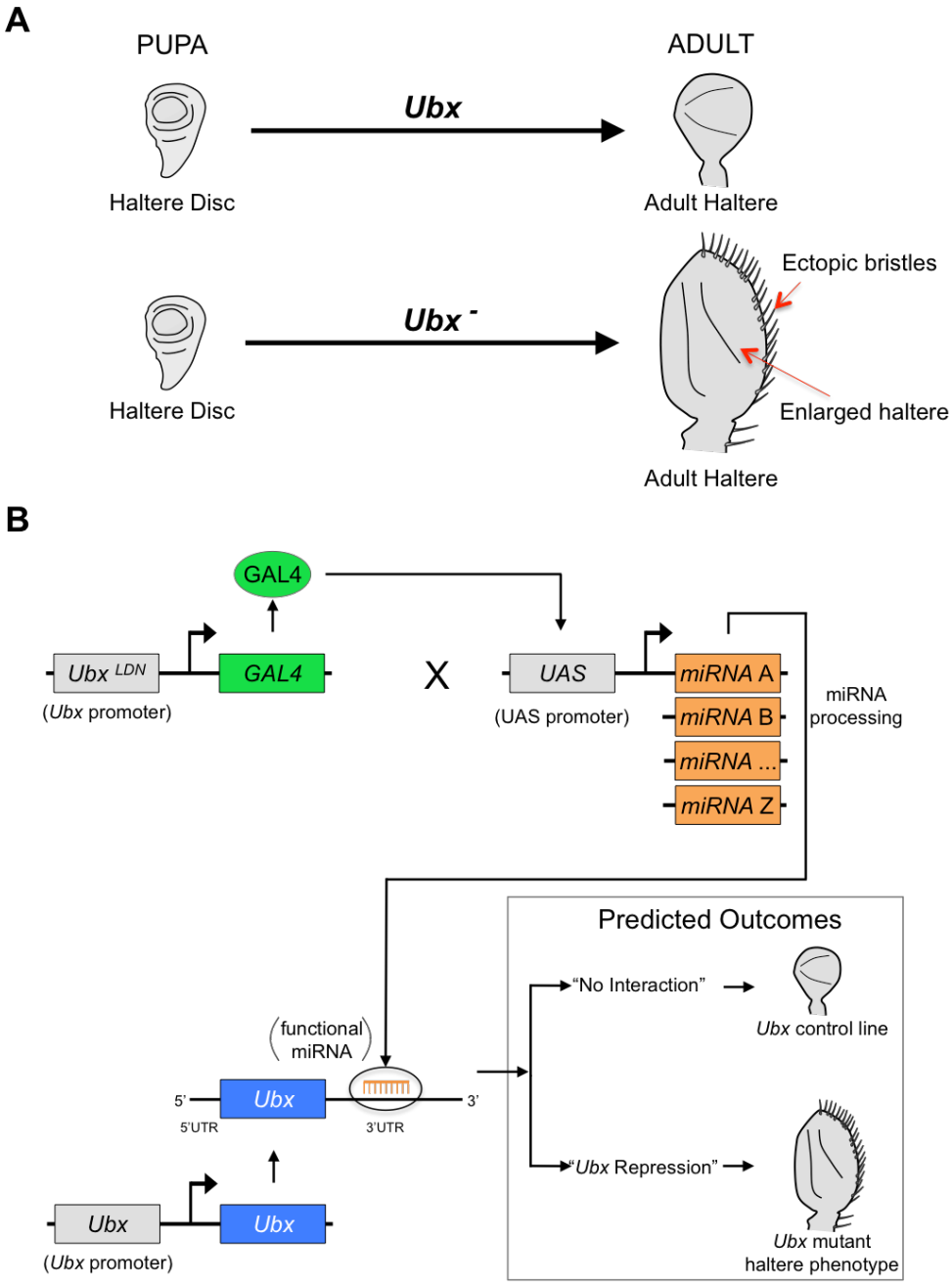


Figure 3.1 Strategy of the Gal4/UAS genetic screen used to identify miRNAs regulating Ubx function *in vivo*.

(Legend on the following page)

Figure 3.1 Strategy of the Gal4/UAS genetic screen used to identify miRNAs regulating Ubx function *in vivo*. (A) The reduction of *Ubx* expression shows the homeotic transformation. Taking advantage of such feature of the haltere, I carried out a genetic screen. (B) Experimental strategy employed in the genetic screen for *Ubx* regulatory miRNAs. With the exception of *bantam*, *miR-281*, *miR-5,4,286* cluster, *miR-6-1,6-2,6-3* cluster and *miR-7*, the rest miRNAs are a collection of UAS-miRNA transgenes with defined insertion in attP2 landing sites including 93 miRNAs and 8 miRNA clusters (Bejarano et al. 2012). I crossed UAS-miRNA flies with the *Ubx*-Gal4^{LDN} driver that mainly expresses in the pouch of the haltere imaginal discs (Figure 4.6). Some of overexpressed miRNAs will target *Ubx* 3' UTR mediating translation repression or mRNA degradation, producing an *Ubx* mutant phenotype. In total, I screened 106 miRNAs and about 28% showed *Ubx* mutant phenotype.

3.2.2 Genetic screen results

In the genetic screen, a substantial amount of miRNA ectopic expressions show morphological changes: fifty-two miRNAs (54% of all individual miRNAs) and five miRNA clusters (50% of all clusters), respectively (Figure 3.8C). This result indicates the haltere as a very sensitive system for the detection of morphological effects induced by miRNA overexpression, and confirms the validity of our approach.

I focused on the miRNAs that induce *Ubx*-like mutant phenotypes. Overall, 28% showed *Ubx*-like mutant homeotic transformations; however, the severity of haltere morphological changes is highly heterogeneous. Among them, 10% out of 28% showed classic haltere to wing transformation that have both bigger halteres and formation of ectopic expressed bristles on the anterior part of the haltere suggesting the repression of *Ubx* by overexpression of these miRNAs (Figure 3.2a, b, c). 18% out of 28% screened miRNAs show mild *Ubx* mutant phenotypes that have either bigger halteres or ectopic expressed bristles indicating other regulations may be involved in the morphological change (Figure 3.2d, e, f, g). Besides targeting *Ubx*, some miRNAs may target *Ubx* downstream genes or other genes that involved in the bristle patterning (Figure 3.2d, e) or cell growth (Figure 3.2f, g).

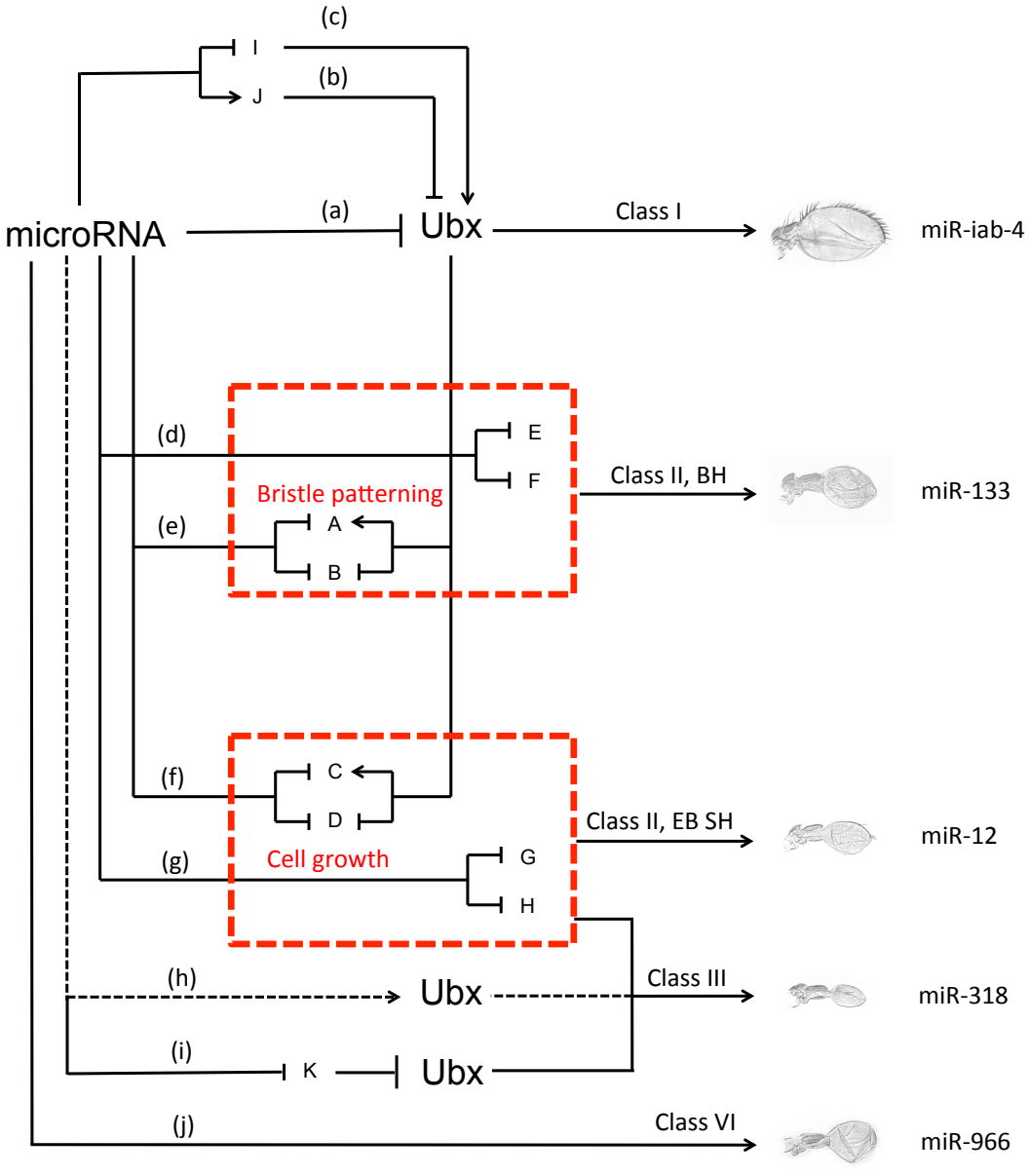


Figure 3.2 miRNAs might act on different components of the genetic program for haltere formation.

A substantial number of overexpressed miRNAs from 96 screened individual miRNAs show haltere morphological changes. 10% of miRNAs show classic haltere-to-wing transformation, which may either regulate *Ubx* directly (a) or activate *Ubx* repressors (b) or silence *Ubx* activators (c). Some miRNAs induce the bigger haltere (BH) without the formation of ectopic bristles (EB) suggesting that besides regulating *Ubx* expression, these miRNAs may also target *Ubx* downstream genes (e) or other genes (d) that are involved in bristle patterning. Conversely, some miRNAs induce the formation of EB with small haltere (SH). Besides regulating *Ubx* expression, these miRNAs may also target *Ubx* downstream genes (f) or other genes (g) that are involved in cell growth. Some miRNAs cause the small haltere, which can be due to different possibilities: (1) these miRNAs target *Ubx* downstream genes (f) or other genes (g) that are involved in cell growth; (2) these miRNAs could also upregulated *Ubx* expression (h); (3) these miRNAs could de-repress *Ubx* (i). About 40% miRNA show no effect to the haltere morphology at all (j). Here I suggest that besides regulating Ubx expression in the haltere, miRNAs may be involved other pathways that control haltere development.

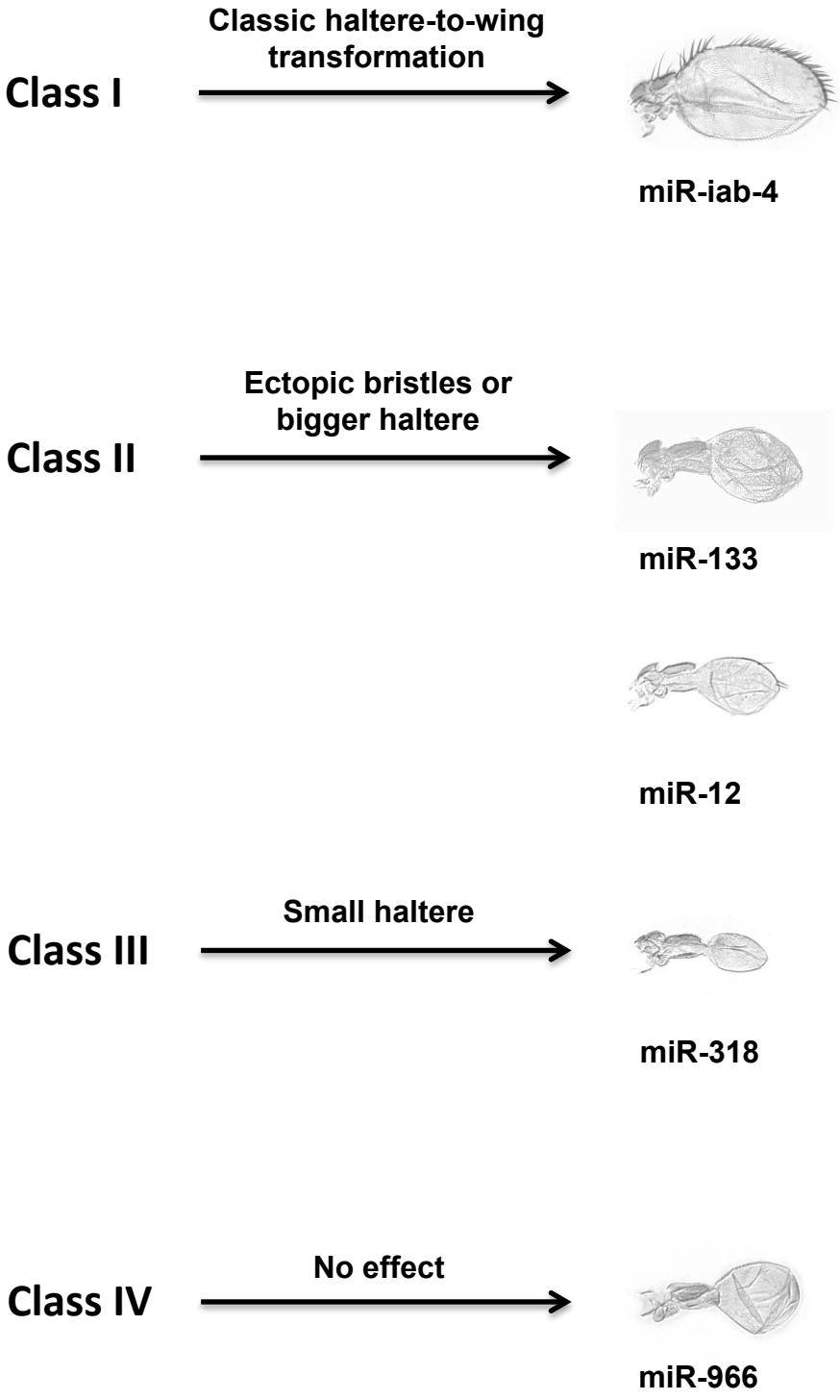


Figure 3.3 A brief summary of miRNA effects on the haltere morphology.

miRNAs in Class I induced classic haltere-to-wing transformation that have bigger halteres and ectopic expressed bristles, such as *miR-iab-4*. Class II miRNAs induced mild *Ubx*-like mutant phenotypes, such as small haltere with the formation of ectopic bristles (*miR-12*), bigger haltere without ectopic bristles (*miR-133*). miRNAs in Class III either showed small haltere or loss of capitellum, I defined this class of miRNAs as other phenotype, like *miR-318*. Class IV miRNAs did not show morphological changes at all with the same GAL4 driver (See Appendix-Table 2).

According to the severity of haltere morphological changes induced by miRNAs, I classified miRNAs into four classes (Figure 3.3). Class I miRNAs induced strong *Ubx*-like mutant phenotype that showed the classic haltere-to-wing transformation. The morphological changes induced by Class II display various phenotypes, such as small haltere with the formation of ectopic bristles, bigger haltere without ectopic bristles, defined as mild *Ubx*-like mutant phenotype. miRNAs in Class III defined as other phenotype showed either small haltere or loss of capitellum. Class IV miRNAs did not produce any morphological changes. I will describe each class in detail in the next sections (Figure 3.4-3.7).

3.2.2.1 Class I (See Chapter 4 for more details)

Overexpression of miRNAs in Class I mimic the strong *Ubx* allele phenotypes which produce classic haltere-to-wing transformations that have bigger halteres and ectopic bristles (Fig 1.4B). Class I includes *miR-iab-4*, *miR-303*, *miR-304*, *miR-92a*, *miR-92b*, *miR-252*, *miR-310*, *miR-311*, *miR-312* and *miR-10* (Figure 3.4D-M). Several groups have reported that *miR-iab-4* regulates *Ubx* expression by either misexpression of *miR-iab-4* in the haltere imaginal discs leads to haltere-to-wing transformation (Ronshaugen *et al*. 2005; Stark et al. 2008; Tyler et al. 2008) and that *miR-iab-4* loss-of-function in *Drosophila melanogaster* causes the increase of *Ubx* expression level in the CNS (Thomsen et al. 2010; Bender 2008). Recent work in our lab has shown that increase of *Ubx* expression level in specific cell lineages in the CNS caused by a *miR-iab-4* loss-of-function allele affects larval behaviour (Picao-Osorio et al. 2015). Another study from our lab has proven that *Ubx* expression level changed in the presence of loss-of-function mutation affecting the *miR-310* cluster and that this led to a phenotypic change in haltere morphology (Richard Kaschula&Claudio Alonso, unpublished). Therefore, these results show that miRNAs in Class I are very likely involved in regulating *Ubx* expression.

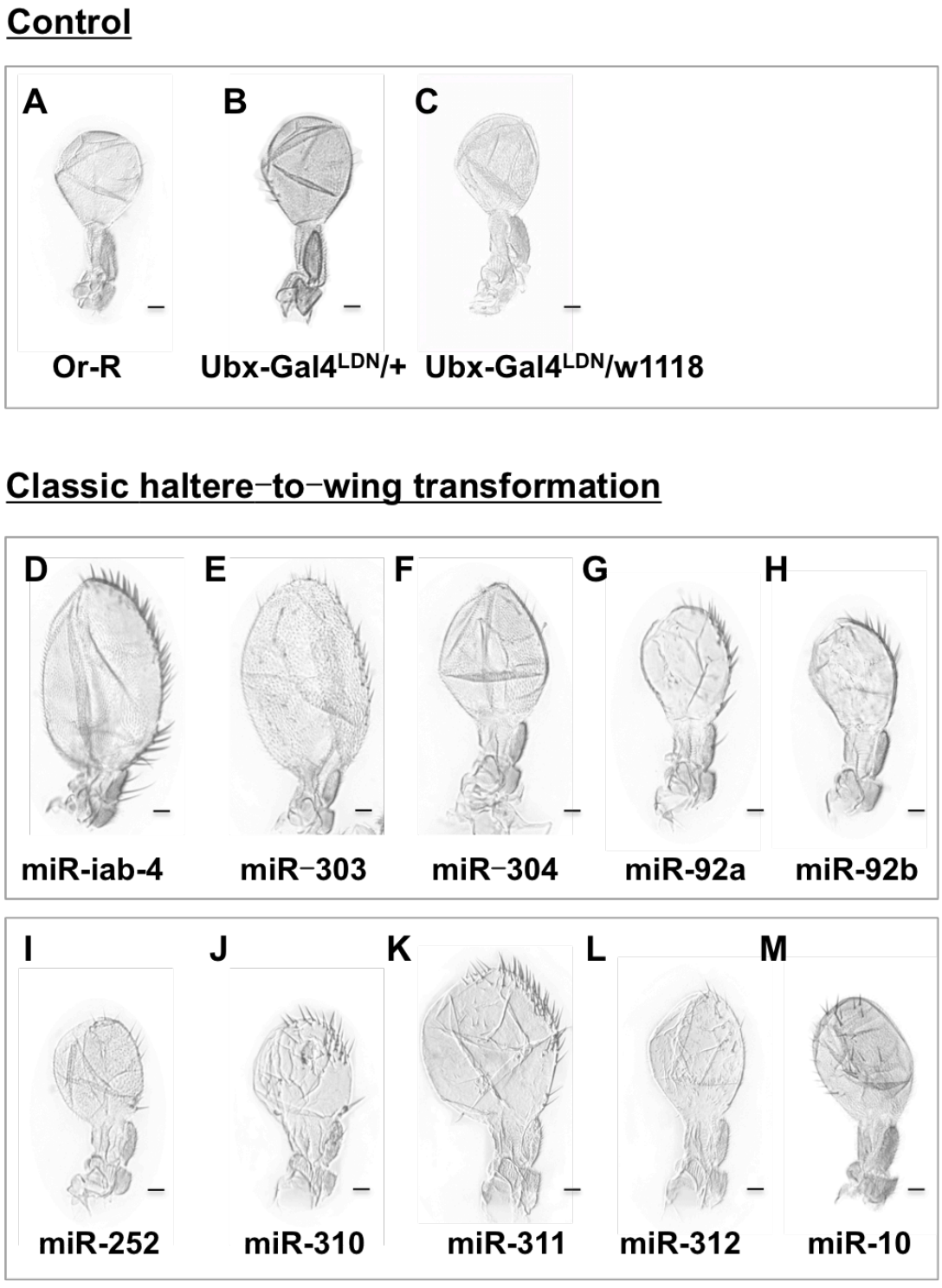


Figure 3.4 Summary of haltere phenotypes caused by misexpression of Class I miRNAs.

All halteres shown are female halteres. (A) Haltere samples of Or-R. (B) Haltere samples from the progeny of *Ubx*-Gal4^{LDN} crossed with Or-R. (C) Haltere samples from the progeny of *Ubx*-Gal4^{LDN} crossed with the *white* gene mutant stock. (D-M) Haltere samples of classic haltere-to-wing transformation induced by different miRNAs. Ten miRNAs, about 10% of all screened miRNAs, induce classic haltere-to-wing homeotic transformation. Scale bar represents 50μm.

3.2.2.2 Class II

miRNAs and miRNA clusters in class II induce various haltere morphological changes that I defined as ‘mild’ *Ubx*-like mutant phenotypes (Figure 3.5D-U). Besides considering those miRNAs in class II regulating *Ubx* expression, they may be involved in other pathways as well.

Although the ectopic expression of *miR-278* and *bantam* both produced bigger halteres and ectopic bristles (Figure 3.5D-E), these effects were not classified as classic haltere-to-wing transformations. Whereas the latter are characterized by a small, flattened haltere that resembles a small wing, and display large ectopic bristles along the anterior region of the haltere, the ectopic expression of the aforementioned miRNAs leads to a spherical haltere morphology and ectopic bristles that are small and cover the whole haltere homogeneously.

Julius Brennecke *et al.* reported that *bantam* controls cell proliferation and apoptosis by targeting the pro-apoptotic gene *hid* (Brennecke et al. 2003). Knud Nairz *et al.* found that misexpression of *miR-278* caused *bantam*-like gain-of-function phenotype (Nairz et al. 2006). Therefore, it is possible that the extra function of the miRNA added to homeotic transformation leads to a different type of morphological change.

The haltere morphological changes induced by *miR-282*, *miR-33* and *miR-12* (Figure 3.5O-Q) all include ectopic bristles, but the size of haltere is similar to the control. Similarly, overexpression of miRNAs *miR-285*, *miR-7*, *miR-8*, *miR-927*, *miR-995*, *miR-263a*, *miR-124*, *miR-79* and *miR-6-1,6-2,6-3*, produced halteres displaying ectopic bristles, but having smaller size (Figure 3.5F-M). Furthermore, the misexpression of *miR-982* led to mild ectopic bristles but smaller halteres (Figure 3.5N). Although different, the morphological effects of the 11 miRNAs and one miRNA cluster described above can be grouped, as all these show the formation of ectopic bristles, but no homeotic transformation for the haltere size. Together, these observations indicate that these miRNAs could potentially regulate *Ubx* expression, but may also control other genes that modulate haltere size.

Here I suggest two possibilities. On the one hand, besides repressing *Ubx* expression, these miRNAs might also be repressing *Ubx* downstream genes that control haltere size (Figure 3.2f). Weatherbee *et al.* reported that different pathways control the development of haltere size and secretion of bristles (Weatherbee et al. 1998). So it is possible that, besides targeting *Ubx*, Class II miRNAs also target *Ubx* downstream genes only influent the haltere size development, but not affecting the formation of ectopic bristles. On the other hand, these miRNAs may also target genes that control haltere size but are not present in the *Ubx* regulatory hierarchy (Figure 3.2g).

Conversely, some miRNAs only induce bigger haltere, but no ectopic bristles. For example, gain-of-function for each of *miR-31b*, *miR-133* and *miR-980* (Figure 3.5R-T) cause bigger haltere but no ectopic bristles compared to the control (Figure 3.5C). I suggest that besides regulating *Ubx* expression, these miRNAs could also target genes downstream of *Ubx* (Figure 3.2e), or other *Ubx* unrelated genes (Figure 3.2d) that control the bristle secretion, leading to a different morphology change. It is possible that the misexpression of miRNA reducing *Ubx* expression level leads to the increase of haltere size and the secretion of ectopic bristles, while other regulations of miRNA suppress the ectopic bristle phenotype. Thus, as an outcome of this additive effect, the size of haltere is increased, but no ectopic bristles are visible.

Cell growth and bristle patterning are under the control of two distinct pathways, and different miRNAs might affect one or the other. It has been shown that miRNAs tend to target different members of the same pathway (Xu & Wong 2008), which is consistent with that obseved for the most cases of the haltere morphological changes by ectopic expression of miRNAs (Figure 3.5F-T).

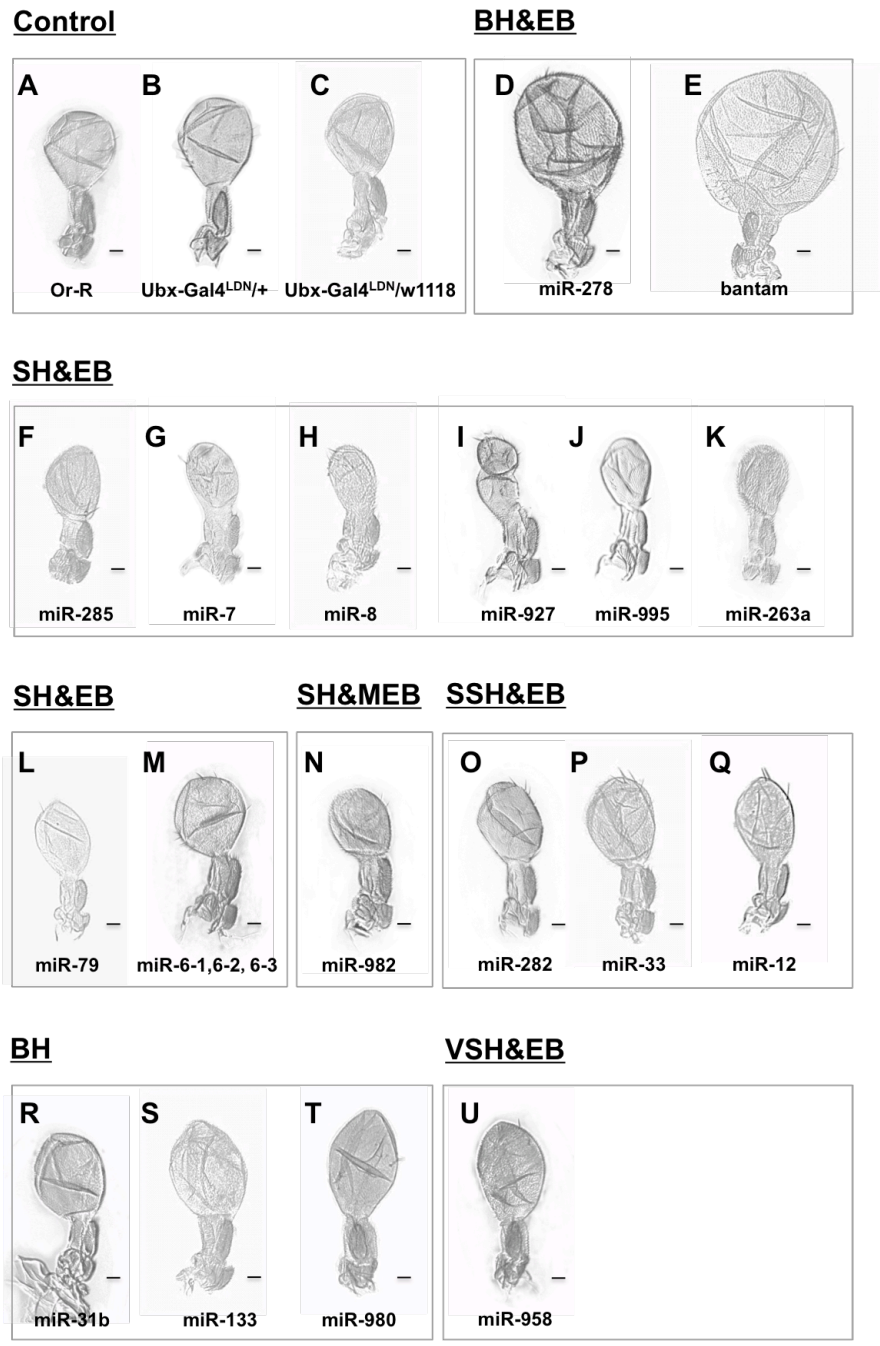


Figure 3.5 Summary of haltere phenotypes caused by misexpression of Class II miRNAs.

All halteres shown are female halteres. (A) Haltere samples of Or-R. (B) Haltere samples from the progeny of *Ubx*-Gal4^{LDN} crossed with Or-R. (C) Haltere samples from the progeny of *Ubx*-Gal4^{LDN} crossed with the *white* gene mutant stock. (D-U) Haltere samples with mild *Ubx* mutant phenotypes induced by different miRNAs. (D-E) Haltere samples with bigger haltere (BH) and ectopic bristles (EB) induced by *miR-278* and bantam. (F-M) Haltere samples with small haltere (SH) and EB induced by different miRNAs. (N) Haltere with SH and mild ectopic bristles (MEB) induced by *miR-982*. (O-Q) Haltere samples with small size haltere (SSH) and EB induced by *miR-282*, *miR-33* and *miR-12*. (R-T) Haltere samples with bigger haltere (BH) induced by *miR-31b*, *miR-133* and *miR-980*. (U) Haltere samples with varied size haltere (VSH) and EB induced by *miR-958*. Here I show 17 miRNAs and one miRNA cluster stimulate mild *Ubx*-like mutant phenotypes. Scale bar represents 50μm.

miR-958 gain-of-function induces the asymmetric size haltere and ectopic bristles (Figure 3.5U). Alqadah *et al*. report that miRNAs are involved in left-right neuronal asymmetry in *C. elegans*, and suggest that this function of miRNA may be evolutionarily conserved across species (Alqadah et al. 2013). Thus, here I suggest that *miR-958* may be involved in the regulation of genes controlling asymmetry development in *Drosophila melanogaster* in addition to the regulation of *Ubx* expression.

3.2.2.3 Class III

All miRNAs and miRNA clusters classified into class III induce smaller halteres (except for *miR-1* and *miR-1012*, in which capitellums almost disappear, Appendix-Figure 1) by misexpression of miRNAs in haltere imaginal discs (Figure 3.6D-DD) compared to controls.

There are several possibilities of how these miRNAs may regulate *Ubx* expression. First, these miRNAs may de-repress *Ubx* expression (Figure 3.2i). For example, Crickmore *et al.* reported that an increase in *Ubx* expression levels leads to the reduction of haltere size (Crickmore et al. 2009). miRNAs in this class may therefore repress the expression of *Ubx* repressor genes, consequently leading to an increased expression of *Ubx* and thus decreasing the haltere size. Second, increasing evidence suggests that miRNA can also upregulate gene expression directly (Vasudevan 2012). The miRNAs in Class III could upregulate *Ubx* expression and cause the decrease of haltere size (Figure 3.2h). Third, it is possible that these miRNAs regulate the *Ubx* downstream genes that control the haltere size (Figure 3.2f). However, it is also possible that these miRNAs regulate other *Ubx* unrelated genes controlling haltere size (Figure 3.2g).

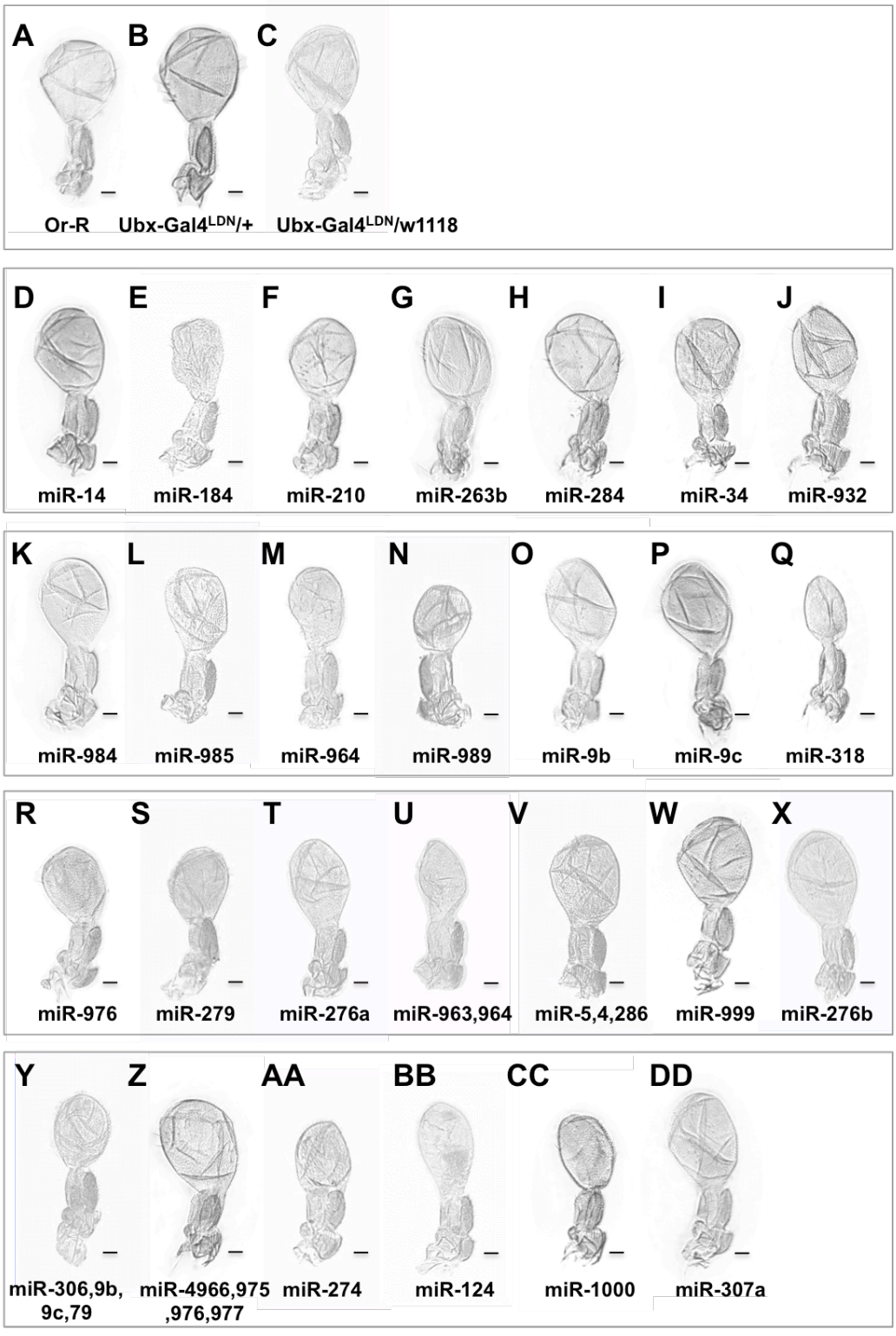


Figure 3.6 Summary of haltere phenotypes caused by misexpression of Class III miRNAs.

All halteres shown are female halteres. (A) Haltere samples of Or-R. (B) Haltere samples from the progeny of *Ubx*-Gal4^{LDN} crossed with Or-R. (C) Haltere samples from the progeny of *Ubx*-Gal4^{LDN} crossed with the *white* gene mutant stock. (D-DD) Haltere samples with small haltere (SH) induced by different miRNAs. Here I show 23 miRNAs and 4 miRNA clusters cause SH by ectopic expression. All pictures were taken under 10x magnification using Leica DMRB microscope. Scale bar represents 50μm.

3.2.2.4 Class IV

miRNAs classified to Class IV (Figure 3.7D-WW) are miRNAs that have no morphological changes at all when they are ectopically expressed in the haltere imaginal discs. For the ectopic expression of individual miRNA, I can confirm that miRNA has no effect on the *Ubx* function.

However, for the miRNA cluster, I can only conclude that the expressions of all the miRNAs together in this cluster have no effects on *Ubx* function. It is possible that some miRNAs in the cluster repress *Ubx* expression, while some upregulate *Ubx* expression. Due to the additive effect, the flies with ectopic expression of all miRNAs in the cluster show halteres similar to controls. For example, *miR-982* and *miR-303* are in the same miRNA cluster and they do not induce haltere morphological changes. But ectopic expression of *miR-303* causes classic haltere-to-wing transformation (Figure 3.4E), while *miR-982* misexpression induces small haltere and mild ectopic bristles (Figure 3.5N). As an outcome of the effect of both miRNAs, the haltere size becomes normal (Figure 3.7TT). However, I cannot explain the loss of the both strong and mild ectopic expressed bristles.

3.2.2.5 Lethal phenotypes

The overexpression of *miR-375*, *let-7* and *miR-310* cluster (Figure 3.8B) using an *Ubx*-Gal4^{LDN} driver led to developmental lethality in the respective progenies. I found, however, that an alternative UAS stock for the *miR-310* cluster (Genotype: +; P {EP}1(2)05510[EP2587]; +) from the Szeged *Drosophila* Stock Center, that when it was overexpressed showed classic haltere-to-wing transformations in a limited number of adult flies. These results support the classification of the *miR-310* cluster into Class I group.

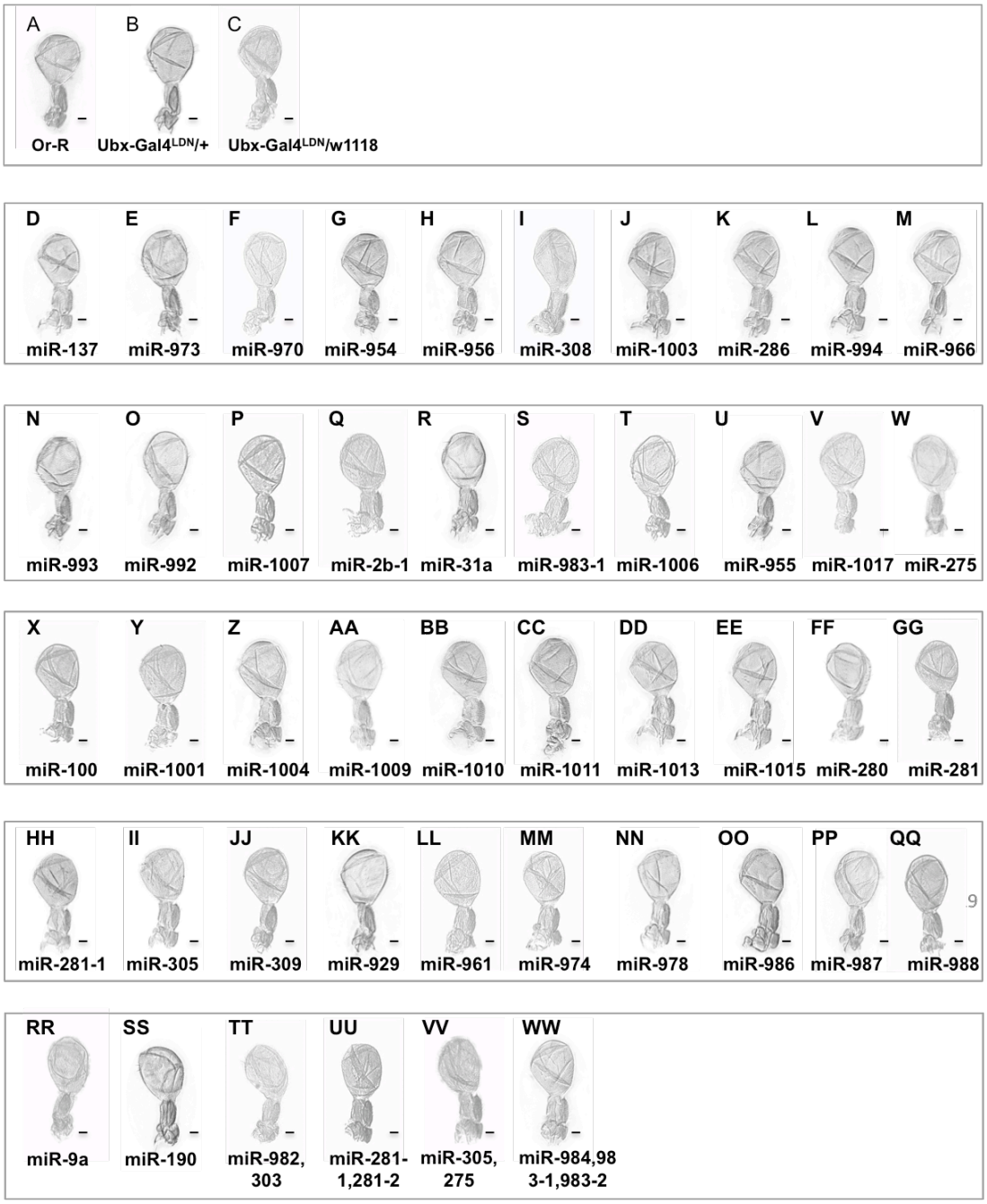


Figure 3.7 Summary of haltere phenotypes caused by misexpression of Class IV miRNAs.

All halteres shown are female halteres. (A) Haltere samples of Or-R. (B) Haltere samples from the progeny of *Ubx*-Gal4^{LDN} crossed with Or-R. (C) Haltere samples from the progeny of *Ubx*-Gal4^{LDN} crossed with the *white* gene mutant stock. (D-WW) Haltere samples from flies misexpressing different miRNAs show no morphological changes. The Class IV group comprises 42 miRNAs and 4 miRNA clusters. Scale bar represents 50μm.

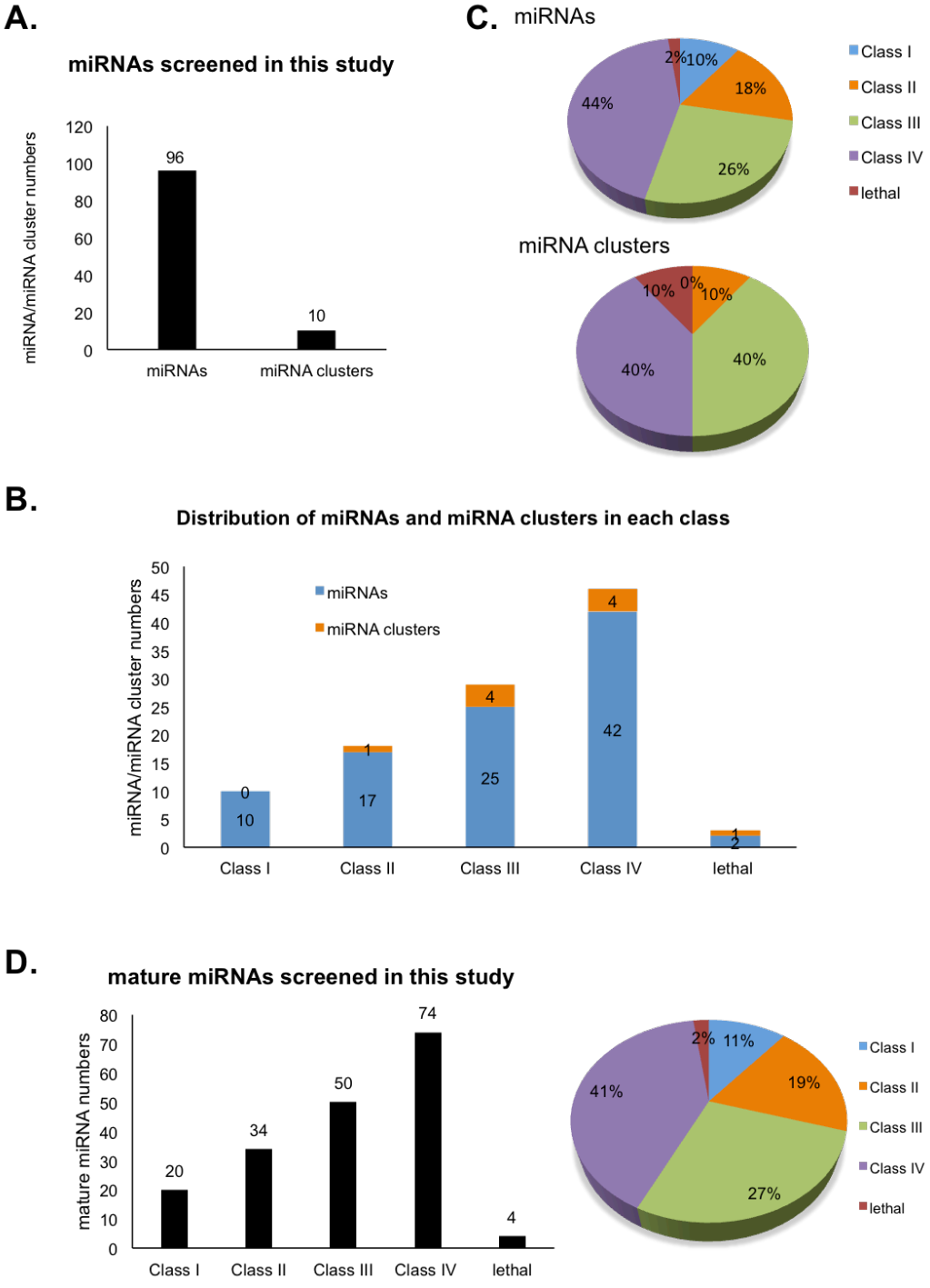


Figure 3.8 Classification and analysis of miRNA genetic screen results.

(A) I screened 96 miRNAs and 10 miRNA clusters in this study (Appendix-Table 2). (B) All miRNAs screened are classified into four classes according to the morphological changes induced by the miRNAs ectopic expression. The miRNA and miRNA cluster number in each class are shown on the bar. (C) The proportion of miRNAs or miRNA clusters in each miRNA class. 28% of the screened miRNAs and 10% of the screened miRNA clusters show *Ubx*-like mutant phenotype. (D) Classification of mature miRNAs screened in this study according to the morphological changes induced by miRNAs. 30% of the mature miRNAs show *Ubx*-like mutant phenotype.

Because the UAS-GAL4 system is a temperature sensitive system, the temperature is another factor that might affect the survival. When I maintained the *miR-375* gain-of-function cross at lower temperature (20°C), 3 males reached adulthood displaying different phenotypes: two males showed bigger halteres and one had smaller halteres. The halteres of these three progenies have lots of bristles attached to the body. Thus, to some degree I could define this miRNA as a mild *Ubx*-like mutant phenotype. The rest of the progeny overexpressing *miR-375* did not enclose and barely survived to pupa stage. Repeating this experiment and having a closer observation of progeny development is needed to ascertain *miR-375* role in the regulation of *Ubx* function.

In summary, I screened 96 miRNAs and 10 miRNA clusters (Figure 3.8A). According to the haltere morphological changes, I separated miRNAs into four classes (Figure 3.8B). 10 miRNAs in Class I showed the classic haltere-to-wing transformation and I defined them as strong *Ubx*-like mutant phenotype miRNAs. Class II included 17 miRNAs and one miRNA cluster that induced mild *Ubx*-like mutant phenotype. 25 miRNAs and 4 miRNA clusters in Class III either showed small haltere or loss of capitellum, I defined this class of miRNAs as other phenotype. Class IV miRNAs were those causing no morphological changes at all after miRNAs gain-of-function with the same Gal4 driver. This class had the largest number of miRNAs, including 42 miRNAs and 4 miRNA clusters. Ectopic expression of two miRNAs and one miRNA cluster were lethal, so they could not be analysed and therefore classified.

From our screen results, 28% of miRNAs and 10% of miRNA clusters (Figure 3.8C) show *Ubx*-like mutant phenotypes, which provides us with a large pool of miRNAs that could potentially regulate *Ubx* expression. Almost half of *Drosophila* miRNAs locate in miRNA clusters (Aravin et al. 2003). However, it has been reported that miRNAs in the same cluster may not be related to each other, and that related miRNAs are not necessary located in the same cluster (Lagos-Quintana et al. 2001; Lau et al. 2001). The miRNA functions in each miRNA cluster may be different, sometimes even having an opposite effect. For example, *miR-982* ectopic expression (Figure 3.5N) and *miR-303* ectopic

expression (Figure 3.4E) show completely different phenotypes. Therefore, instead of taking the risk of different miRNAs having a different function in the same cluster, I only isolated and considered individual miRNAs screened for further analysis.

The mature miRNA is produced from the pre-miRNA and regulates mRNA expression. According to the miRNA update from miRBase 21, I performed an in depth analysis of the mature miRNAs. In the genetic screen, I screened 182 mature miRNAs (Figure 3.8D) and those with *Ubx*-like mutant phenotypes accounted for 30% of all mature miRNAs screened.

3.2.3 miRNA expression levels in pre-pupal haltere imaginal discs do not correlate with the strength of haltere phenotypes.

From all 96 screened miRNAs, 28% of miRNA transgenes show *Ubx*-like mutant phenotype by ectopic expression. miRNA overexpression may exceed the native expression level of that miRNA and lead to unspecific targeting. The haltere morphological change can occur due to the biological function of the miRNAs, but it also can be the result of unspecific targeting caused by high level of miRNA expression in the developing haltere. In order to rule out the latter potentially “false positive” result, it is necessary to check the native expression level of miRNAs. Thus, I can determine if there is any correlation between the miRNA expression level and haltere phenotypic changes.

I compared the miRNA expression levels in white pre-pupal haltere imaginal discs between different miRNA classes. The miRNA expression level in haltere was acquired from RNA-seq data obtained from white pre-pupal (WPP) haltere imaginal discs (Figure 3.10A). In addition, Figure 3.9 illustrates the miRNA expression levels in different miRNA classes.

The miRNA expression levels between different miRNA classes are similar with a few exceptions. In Class I, the expression level of *miR-92b-3p* is extremely high compared to the other miRNAs in the same class (Figure 3.9). However,

the haltere morphological change caused by *miR-92b* is not the most severe one across 10 miRNAs in Class I (Figure 3.4). So it is very unlikely that the high expression level of *miR-92b-3p* induces the phenotypic change. Furthermore, the *miR-8-3p* in Class II, *miR-9c-5p* in Class III and *miR-9a-5p* (1.69E+15 FPKM) in Class IV are the most highly expressed miRNAs in each Class, respectively (Figure 3.9). Nevertheless, apart from *miR-8* shows mild *Ubx*-like mutant phenotype (Figure 3.5H), *miR-9a* (8.68E+14 FPKM) (Figure 3.7RR) and *miR-9c* (1.08E+15 FPKM) (Figure 3.6P) do not induce any phenotype, although their expression levels are much higher than *miR-8* (2.69E+14 FPKM). This result suggests that the *Ubx*-like mutant phenotype of haltere in the genetic screen is not due to the unspecific targeting caused by the saturation of miRNA expression level.

To have a statistically meaningful comparison for miRNA expression in different phenotypic classes, I performed a statistical analysis by using Prism (one-way ANOVA, nonparametric test, Kruskal-Wallis test). I found there were no significant differences of miRNA expression levels among different miRNA classes (Figure 3.9). Thus, *Ubx* mutant phenotype is likely due to the biological function of the miRNAs instead of the result of unspecific targeting caused by high miRNA expression level in the developing haltere, and there is no correlation between the miRNA expression levels and haltere phenotypic change.

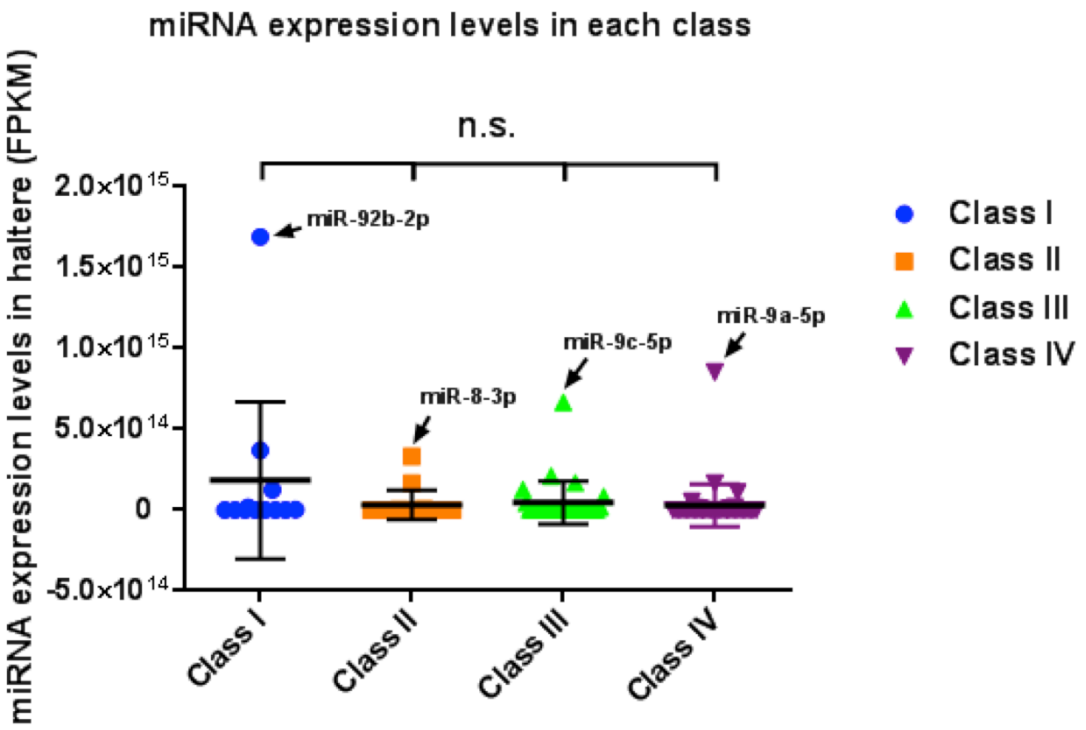


Figure 3.9 Correlation between miRNA expression levels and phenotypes caused by ectopic expression of miRNAs.

Here I show the expression levels of miRNAs in different phenotypic classes measured by RNA-seq using the samples of pre-pupal haltere imaginal discs **(collected by our previous lab member Dr. Richard Kaschula)**. The miRNA expression level was calculated by FPKM (fragments per kilobase of transcript per million fragments mapped). There are no significant differences among miRNA expression levels from different classes (one-way ANOVA, nonparametric test, Kruskal-Wallis test) statistic test (p-values>0.05; non-significant; n.s.). Thus, there is no correlation between the miRNA expression levels and haltere phenotypic changes. The most highly expressed miRNAs in each Class are marked.

3.2.4 Validation of RNA-seq results.

RNA-seq, a next generation sequencing (NGS) technology for deep sequencing of RNA, can be used to evaluate and quantify the RNA expression in a specific cell or tissue of an organism, including low abundant RNAs (Mortazavi et al. 2008; Wilhelm et al. 2008). Thus, it is a commonly used high-throughput technique to sequence the miRNA transcriptome. However, the precise expression levels of miRNAs still require validation using other approaches due to limitations in sensitivity and specificity of RNA-Seq. Stem-loop RT followed by TaqMan PCR analysis is a highly sensitive, specific and precise method to

quantify mature miRNAs with no genomic DNA contamination (Chen et al. 2005). This method is a good complement to RNA-Seq and can be used to validate the miRNA expression.

To validate the reliability of our RNA-seq result, I selected some miRNAs with different expression levels (the miRNAs are defined as 'none', 'low', 'medium', 'high' and 'very high' according to the abundance of miRNAs in haltere by RNA-seq data, shown as Appendix-Table 1) from different miRNA phenotypic classes (Figure 3.10B). Then I confirmed the expression of these miRNAs by applying TaqMan RT-PCR to quantify the miRNA expression in the white pre-pupal haltere imaginal discs using the same RNA samples used to make the RNA-seq library (Figure 3.10A). The results were normalized to U27 and shown in Figure 3.10C.

Comparing the patterns of miRNA expression between Figure 3.10B and Figure 3.10C, I found that the relative miRNA expression levels within different phenotypic classes are very similar. The miRNA expression profiles in Class I, Class II, Class III and in miRNAs that caused lethal effect are very similar based on both miRNA-seq and TaqMan RT-PCR results. For example, in phenotypic Class I, *miR-92a-3p* (1.25E+14 FPKM) has the highest expression by RNA-seq. The second and third ones are *miR-252-5p* (6.67E+11 FPKM) and *miR-304-5p* (66412300000 FPKM), respectively. *miR-iab-4-5p* (9387530000 FPKM) has the lowest expression in this class. The order of miRNA according to abundance in this class is the same based on the results of TaqMan RT-PCR.

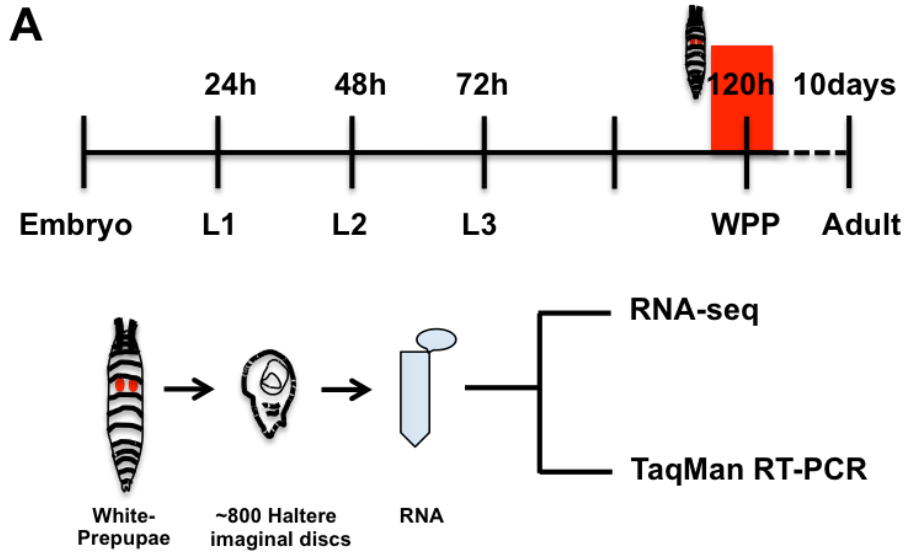
However, the miRNA expression pattern in Class IV measured by TaqMan RT-PCR was different from miRNA sequencing results. From the result of RNA-seq, *miR-137-3p* (65712700000 FPKM), *miR-190-5p* (2.12E+11 FPKM) and *miR-9a-5p* (8.53E+14 FPKM) are all more abundant than *miR-1006-3p* (39847400000 FPKM), while *miR-1006-3p* (0.007816) has higher expression level when compared to *miR-137-3p* (0.001642) and *miR-190-5p* (0.003381) by TaqMan RT-PCR.

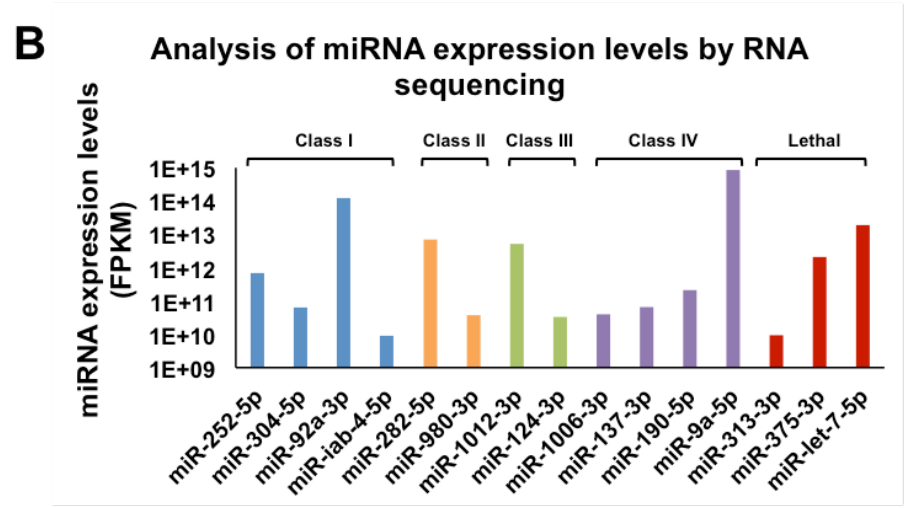
Also, the relative miRNA expressions between different phenotypic classes are

very similar, except for *miR-980-3p*, *miR-124-3p* and *miR-1006-3p*. For example, the expression of *miR-980-3p* (9.92E-05) is only higher than *miR-iab-4-5p* (9.05E-05) by TaqMan RT-PCR. Whereas, by RNA-seq, the expression of *miR-980-3p* (37550100000 FPKM) is much higher compared to *miR-124-3p* (33206200000 FPKM), *miR-iab-4-5p* (9387530000 FPKM) and *miR-313-3p* (9387530000 FPKM).

Although with exceptions described above, the high similarity of the relative miRNA expression patterns between RNA-seq and TaqMan RT-PCR indicates the reliability of the RNA-seq data. I suggest that the RNA-seq data is suitable to be applied for comparing the miRNA expressions between different phenotypic classes.

Considering the high cost of the TaqMan RT-PCR, I only selectively tested some miRNAs to re-examine the correlation between the miRNA expression levels and haltere phenotypic changes. Comparing miRNA expression by using the TaqMan RT-PCR data, I found that miRNA expression levels are quite similar among different classes with some exceptions. The expression levels of *miR-92a-3p* (2.037312) in Class I and *miR-9a-5p* (15.83335) in Class IV are extremely high compared to the other tested miRNAs and there is statistical significant difference between the expression of these two miRNAs and the rest of the miRNA. The expression levels of *miR-iab-4-5p* (9.05E-05) in Class I, *miR-980-3p* (9.92E-05) in Class II and *miR-313-3p* (0.00021) with lethal phenotype are relatively lower and there is statistical significant difference between the expression of these two miRNAs and the rest of the miRNA. Both high and low miRNA abundance distribution in phenotypic Class I and other phenotypic classes by two different techniques prove that there is no correlation between the miRNA expression levels and haltere phenotypic changes. Thus, the morphological change induced by ectopic miRNA expression appears to be related to the miRNA biological function rather than an artifact caused by elevated level of miRNAs produced in the gain-of-function screen.





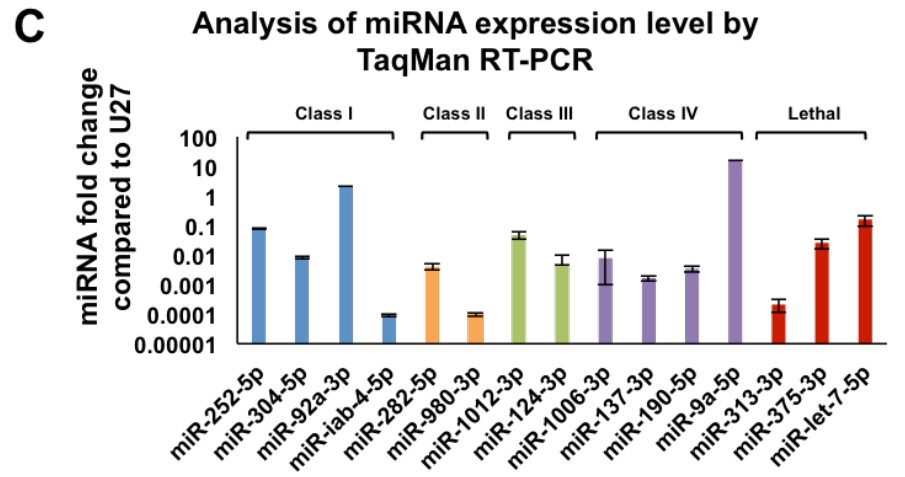


Figure 3.10 Analysis of miRNA expression levels in pre-pupal haltere imaginal discs by RNA-seq and TaqMan RT-PCR.

(Legend on the following page)

Figure 3.10 Analysis of miRNA expression levels in white pre-pupal haltere imaginal discs by RNA-seq and TaqMan RT-PCR. (A) Only white pre-pupae (WPP, developmental stage that lasts about 1h) were selected for dissection of haltere imaginal discs to make the tissue populations as homogenous as possible. Around 800 haltere imaginal discs were dissected and applied for RNA extraction. The sample was separated into two portions and used for RNA-seq and TaqMan RT-PCR, respectively. **The sample collection was done by Dr. Richard Kaschula and Dr. Ana Bomtorin.** (B) Graph shows the various miRNA expression levels of different phenotypic classes measured by RNA-seq **(collected by our previous lab member Dr. Richard Kaschula)**. The blue columns, orange columns, green columns, purple columns refer to expression levels of miRNAs from Class I, Class II, Class III and Class IV, respectively, and the red columns stand for expression levels of miRNAs that ectopic expression cause the lethal effect. (C) Graph denotes miRNA expression levels in white pre-pupae haltere imaginal discs measured by TaqMan RT-PCR. The miRNA expression levels are normalized to U27, an artificial TaqMan miRNA control assay. The colors meanings of each columns are indicated the same as in B. This experiment has three repeats. Here I show that the relative miRNA expression level by each technique is quite similar, which indicated the reliability of RNA-seq data. Also, comparing the miRNA expression across four miRNA classes by both RNA-seq and TaqMan RT-PCR, I confirm that there is no correlation between the miRNA expression levels and haltere phenotypic changes.

3.2.5 Investigating miRNA inputs on the developmental gene *Ubx* using ten bioinformatic prediction programs

In this study, I screened 96 miRNAs and 10 miRNA clusters. Based on a phenotypic analysis of adult halteres, about 28% miRNAs can potentially regulate *Ubx* expression. When compared to this experimental process, bioinformatic approaches are a much more convenient way to find miRNAs that target a particular 3' UTR. However, different bioinformatic prediction tools have varying degrees of accuracy.

To examine the performance of current bioinformatic prediction programs, I performed several parallel bioinformatic predictions for *Ubx* using all ten miRNA targeting prediction programs available for *Drosophila*, using as an input the miRNAs presented in the genetic screen. The settings used for all softwares are described in detail in Chapter 2. The results of the predictions are shown in Figure 3.11.

Among all four programs using local predictions, RNAhybrid returned the largest number of predicted targets with all 182 mature miRNAs predicted to target *Ubx*. The second largest one was PITA that got 114 out of 182 miRNAs predicted to

target *Ubx*. The predicted targets of miRanda and RNA22 were low: 44 and 36 mature miRNAs were predicted to target *Ubx* by miRanda and RNA22, respectively.

For those that have more limited miRNA input, such as ComiR, EMBL, DIANA, miRecord, PicTar and TargetScan, miRecord programs identified more limited number of target predictions: 51 out of 96 miRNAs were predicted to target *Ubx*. ComiR predicts that 35 out of 181 miRNAs target *Ubx*, whereas 6 out of 39, 7 out of 46, 22 out of 97 and 19 out of 181 miRNAs were predicted to target *Ubx* by the EMBL, PicTar, TargetScan, and DIANA miRNA targeting prediction tools, respectively. The screened miRNA showed great potential to regulation Ubx expression according to bioinformatic predictions.

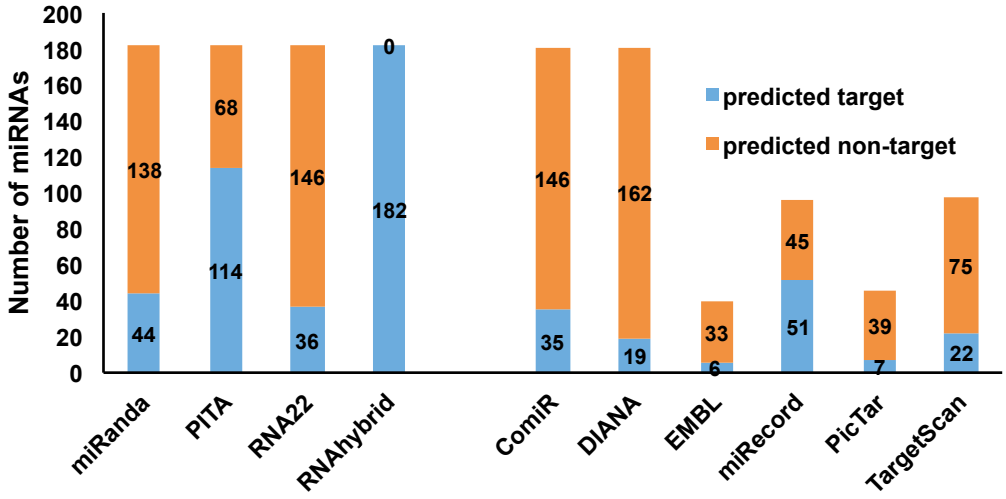


Figure 3.11 miRNA target predictions for *Ubx* by using ten bioinformatic programs.

Comparison of miRNA target-prediction results for ten bioinformatic programs. The y-axis indicates the number of mature miRNAs predicted to target *Ubx*. Each column is composed of two parts: predicted targets and predicted non-targets. Blue bars refer to the number of miRNAs that were predicted to target *Ubx* 3' UTRs, including EMBL, miRanda, PicTar, PITA, TargetScan, RNAhybrid, RNA22 and ComiR. In the case of DIANA, blue bars refer to the number of miRNAs predicted to target both *Ubx* CDSs and 3' UTRs. The orange bars indicate the number of miRNAs that were inputted in the computational prediction but not predicted to target *Ubx*. The x-axis lists the different bioinformatic miRNA targeting prediction tools used in these analyses. RNAhybrid returns the largest number of miRNAs predicted to target *Ubx*.

3.3 Discussion

In this chapter, I explore the question: what are the potential *Hox* gene *Ubx* regulatory miRNAs in the context of development? I carried out the genetic screen on *Drosophila* haltere. The *Drosophila* haltere is a very sensitive system to detect the *Ubx* expression level changes, which shows a haltere to wing transformation when *Ubx* expression is decreased (Lewis 1978). In contrast, haltere becomes smaller when *Ubx* expression is increased (Crickmore et al. 2009; Ronshaugen et al. 2005). By using the UAS-GAL4 system, I ectopically expressed miRNAs by *Ubx*-Gal4^{LDN}. *Ubx*-Gal4^{LDN} driver is a very good tool for genetic screen for the following reasons: 1) *Ubx*-Gal4^{LDN} is not expressed during embryogenesis, which allows *Drosophila* to develop to later stages even if miRNA overexpression causes a deleterious effect on *Drosophila* development (Navas et al. 2006). This strategy allows the survival of a high proportion of adult flies expressing miRNA ectopically. 2) The expression domain of *Ubx*-Gal4^{LDN} is partially overlapped with *Ubx* in the haltere, which provides a good system to study the changes of *Ubx* expression level by overexpressing miRNA. 3) *Ubx*-Gal4^{LDN} is a hypomorphic allele of *Ubx* showing an *Ubx* mutant phenotype. This system will increase the visibility of the effect caused by even a mild repression of Ubx expression by miRNA.

Most evidence of miRNA-mediated *Hox* gene regulation during animal development comes from the miRNAs encoded in *Hox* cluster. In vertebrates, only *miR-10* and *miR-196* are experimentally proven to regulate *Hox* gene expression (Yekta et al. 2004; He et al. 2012; Woltering & Durston 2008; McGlinn et al. 2009). For example, *miR-196* direct cleavage of *HOXB8* leads to the repression of this *Hox* gene expression in mouse embryos (Yekta et al. 2004). In *Drosophila*, it has been shown that *miR-iab-4* and *miR-iab-8* regulate *Ubx* expression (Bender 2008; Garaulet et al. 2014; Gummalla et al. 2014; Ronshaugen et al. 2005; Stark et al. 2008; Thomsen et al. 2010; Tyler et al. 2008).

In this genetic screen I found many other miRNAs that are not encoded in *Hox*

cluster could potentially regulate *Hox* gene *Ubx* expression (Figure 3.12). For example, miRNAs in *miR-310* cluster are located in the second chromosome. In total, I screened 96 miRNAs and 10 miRNA clusters transgenes. Including *Hox* cluster miRNAs (*miR-iab-4* and *miR-10*), 28% of miRNAs (27 miRNAs) and 10% of miRNA clusters (one miRNA cluster) screened could potentially regulate *Hox* gene *Ubx* expression. Thus, I showed that *Hox* genes might not be only specifically regulated by *Hox* cluster miRNAs. Conversely, *Hox* cluster miRNAs appears not to be necessary to regulate *Hox* genes. For example, *miR-993* is encoded in *Hox* cluster, but ectopic expression of this miRNA shows no phenotype at all (Figure 3.7N).

According to the haltere morphological change by ectopic expression of miRNAs, I classified miRNA into four classes: Class I, strong *Ubx*-like mutant phenotype; Class II, mild *Ubx*-like mutant phenotype; Class III, other phenotype. Class IV, no phenotype.

Except for the classic haltere-to-wing transformation, I have classified some morphological abnormal flies as mild *Ubx*-like mutant phenotype. This category comprises miRNAs that induces the formation of bigger halteres or ectopic bristles. It is possible that these candidate miRNAs may not interact with *Ubx* directly, instead, they may actually target *Ubx* downstream genes, such as *Salr*, *blistered* or *AS-C* (Weatherbee et al. 1998) or other genes leading to the change of haltere size or formation of ectopic bristles. Further experiments such as immunohistochemistry can be applied to test whether *Ubx* expression level changes by ectopic expression of the miRNAs in Class II.

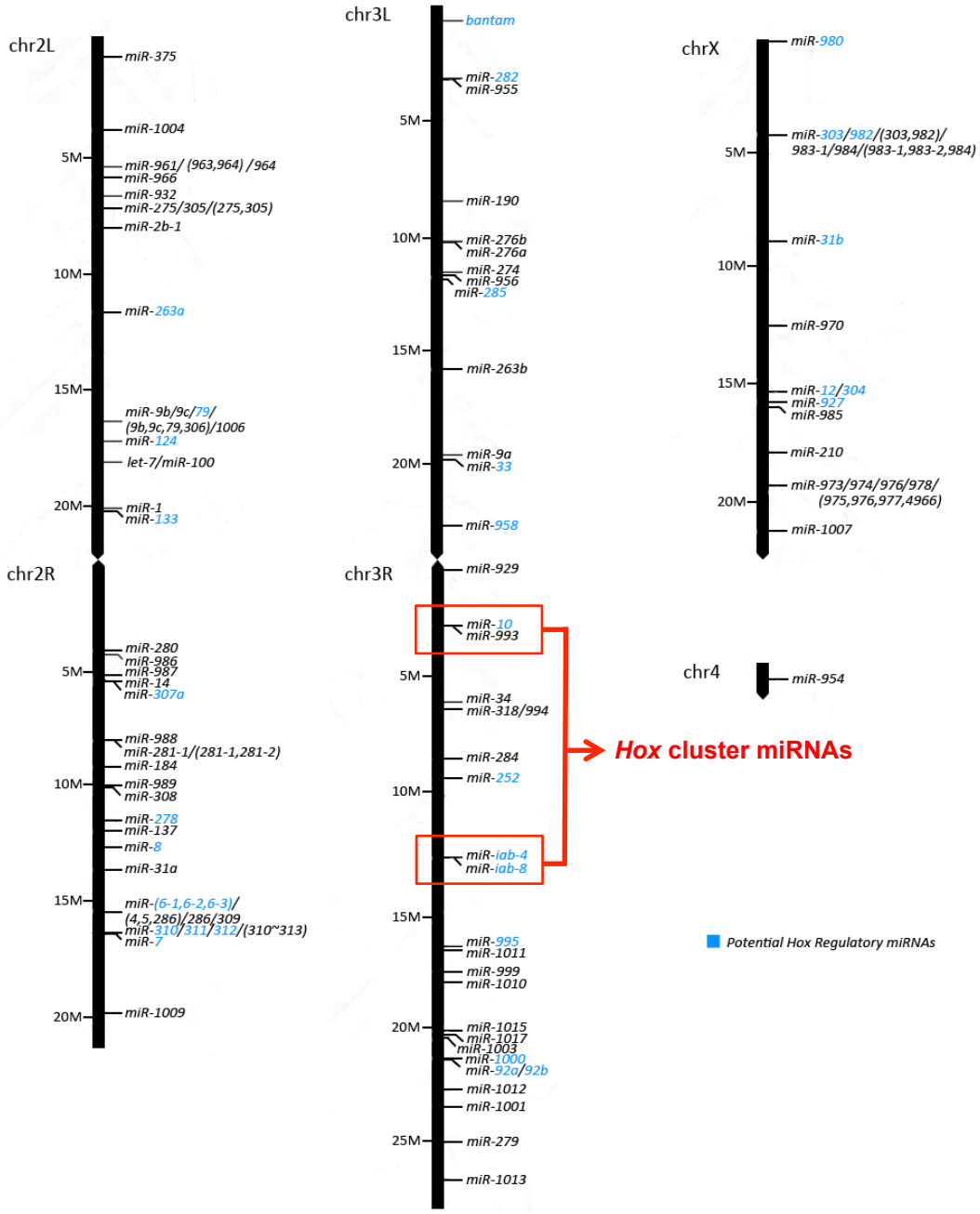


Figure 3.12 The genomic landscape of screened miRNA genes.

The potential *Ubx* regulatory miRNAs are marked by blue. I find 28% of miRNAs (27 miRNAs) and 10% of miRNA clusters (one miRNA clusters) screened have the potential to regulate *Hox* gene *Ubx* expression. The miRNA encoded in the *Hox* cluster is not necessary to regulate *Hox* gene expression such as *miR-993*, while about 24 miRNAs and one miRNA cluster located outside the *Hox* cluster could regulate *Hox* gene *Ubx* expression.

Ectopic expression of miRNAs in Class III causes small haltere or loss of capitellum. It has been shown that increased *Ubx* expression level in haltere to certain extent can reduce haltere size (Smolik-Utlaut 1990; Crickmore et al. 2009). Therefore, I suggest that miRNA in Class III may induce the upregulation of *Ubx* expression or *Ubx* downstream targets. Recently, increasing evidence shows that miRNAs upregulate or de-repress gene expression (Vasudevan 2012). However, the molecular mechanism behind this type of regulation is unclear.

Almost half of *Drosophila* miRNAs are located in miRNA clusters (Aravin et al. 2003). The miRNAs in the same miRNA cluster could have same or related function. For example, ectopic expression of *miR-305* and *miR-275* induce the same phenotype in haltere (Figure 3.7W, II). However, it has been reported that miRNAs in the same cluster may not be related to each other (Lagos-Quintana et al. 2001; Lau et al. 2001). The miRNA functions in each miRNA cluster may be different, sometimes even having an opposite function. For example, *miR-982* ectopic expression (Figure 3.5N) and *miR-303* ectopic expression (Figure 3.4E) show completely different phenotypes, despite being located in the same cluster.

Apart from *Ubx* function study, this genetic screen also opens the possibility for other functional study of miRNAs. For example, ectopic expression of *miR-278* and *bantam* do not cause traditional haltere-to-wing transformations. *miR-278* and *bantam* overexpression induce dramatic increase of the haltere size and ectopic bristles. This phenotype is consistent with previous studies that *bantam* and *miR-278* control cell proliferation and apoptosis (Brennecke et al. 2003; Nairz et al. 2006).

In the end of this chapter, by using the RNA-seq data and the TaqMan RT-PCR validation, I found that there seems no correlation between the strength of haltere phenotypes and the endogenous miRNA expression level in wild-type white pre-pupal haltere imaginal discs. Thus, it seems that the haltere phenotypic changes are qualitative miRNA effects and there is no bias in our genetic screen. In addition, the screened miRNA showed great potential to

regulation Ubx expression according to bioinformatic predictions. I am going to validate the genetic screen result and examine the bioinformatic prediction programs in the next chapter.

In summary, I describe a genome-wide screen of 96 miRNAs and 10 miRNA clusters, and find that 28% miRNAs and 10% miRNA clusters could induce *Ubx* mutant phenocopies. This strategy has provided a significant number of potential *Hox* regulatory miRNAs in a development context.

CHAPTER 4 The molecular basis of miRNA-induced haltere phenotypes and their relation to *Ubx* regulation.

4.1 Introduction

In the previous chapter, I described the results of a genetic screen for 96 miRNAs and showed that about 10% (10 miRNAs) of all screened miRNA transgenes show a classic haltere-to-wing transformation. Due to the simplicity of the genetic cross strategy and its visible phenotypic readout, the aforementioned genetic screen is a simple and efficient approach to identify miRNA regulation of *Ubx* mRNAs. However, one potential pitfall of this approach is that miRNAs may regulate *Ubx* downstream targets and induce a *Ubx*-like mutant morphological change or otherwise affect genes related to haltere size and bristle development that are not under the control of *Ubx* regulatory hierarchies, thus causing a similar phenotype. Therefore, further validation of genetic screen results using other approaches is crucial to confirm genuine miRNA regulation of *Ubx*. Immunohistochemistry provides a simple and efficient way to evaluate changes in specific protein expression while providing spatial information on such fluctuations. In the present chapter, I used this technique to address the following question: what are the effects of miRNAs on *Ubx* expression?

Another limitation of our genetic screen is that although miRNAs may affect *Ubx* expressions, such changes could occur as a result of both direct and indirect targeting. Do the candidate miRNAs physically interact with *Ubx* 3' UTRs? An *in vivo* approach to test the direct interaction between miRNAs and *Ubx* mRNAs can be performed using fluorescent *Ubx* 3' UTR reporters expressed in the developing haltere imaginal disc. A combined expression of UAS-miRNA and UAS-3' UTR transgenes, using a specific GAL4 driver enables both constructs to be co-expressed in certain tissue, which is a valuable approach to test the tissue specific direct interaction between miRNAs and mRNAs (Duffy 2002).

Bioinformatic prediction, compared to more demanding experimental

approaches, have emerged as a complementary and potentially faster approach to identify potential sites for miRNA-mRNA interactions. Since 2003, several groups have used bioinformatic approaches to generate algorithms that predict miRNA targets (Stark et al. 2003; Enright et al. 2003; Rajewsky 2006; Kiriakidou et al. 2004). Among them are DIANA, EMBL, miRanda, PicTar, PITA, RNA22, TargetScan, RNAhybrid, ComiR and miRecord (Grün et al. 2005; Stark et al. 2003; Kheradpour et al. 2007; Saito & Sætrom 2010; Enright et al. 2003; Sethupathy et al. 2006), programs often used for miRNA target prediction in *Drosophila melanogaster*.

In this chapter, I have focused on the analysis of miRNAs in Class I (see previous chapter). Firstly, I have used immunohistochemistry to confirm that nine miRNAs in Class I do regulate *Ubx* expression in the haltere imaginal discs. Secondly, I have employed a genetic approach to address the question of how many miRNAs in Class I directly interact with *Ubx* 3' UTRs. To this end, UAS-*Ubx* 3' UTR and UAS-miRNA genetic constructs were co-expressed in haltere imaginal discs using an *scalloped*-GAL4 (*sd*-GAL4) driver, revealling six miRNAs in Class I directly interact with *Ubx* 3' UTR. This provides a high-confidence list of potential *Ubx*-regulatory miRNAs. Lastly in this chapter, miRNAs that directly interact with *Ubx* 3' UTR are used to analyse the genuine target sites by using different bioinformatic prediction programs.

4.2 Results

4.2.1 Phenotypic analysis of miRNAs from Class I using adult haltere cuticle preparations.

miRNAs in phenotypic Class I show classic haltere-to-wing transformations with subtle variations. To fully charaterise these small variations, I used adult haltere cuticle preparations for halteres and quantified haltere phenotypes at higher magnification (Figure 4.1).

Ubx-Gal4^{LDN} is a GAL4 insertion in heterozygosis produces a mild *Ubx*-like

mutant phenotype (Navas et al. 2006), with 96.5% (55/57) of the progeny showing slightly bigger halteres (Figure 4.1B, B') and a very small number (2 out of 57) of these progenies having one longer, thicker bristle on the proximal of the capitellum (Appendix-Table 2).

Ectopic expression of *miR-iab-4* results in the most obvious haltere-to-wing transformation (Figure 4.1C, C'), consistent with previous reports (Ronshaugen et al. 2005). The haltere that results from overexpression of *miR-iab-4* is flattened and similar to the shape of a reduced wing. The size of the haltere by ectopic expression of *miR-iab-4* (Figure 4.1C) seems six times that of the control (Figure 4.1B). The anterior part of the haltere (Figure 4.1C') developed several rows of long, thick bristles similar to those on the wing margin (Bejarano et al. 2012). The homeotic transformation induced by *miR-iab-4* in the screening is in perfect agreement with previous studies showing that *miR-iab-4* regulates *Ubx* expression in *Drosophila* (Bender 2008; Garaulet et al. 2014; Gummalla et al. 2014; Ronshaugen et al. 2005; Stark et al. 2008; Thomsen et al. 2010; Tyler et al. 2008).

The haltere with ectopic expression of *miR-303* (Figure 4.1D) looks to have similar shape and size to the one with *miR-iab-4* overexpression, but with fewer and thinner ectopic bristles on the anterior part of the haltere (Figure 4.1D'). Haltere with ectopic expression of *miR-304* (Figure 4.1E, E') seems to have a similar type of ectopic bristles as that with ectopic expression of *miR-303*, but their size looks one-third smaller to that of *miR-303*. *miR-252* ectopic expression haltere (Figure 4.1H) looks to have increased size compared to wild-type control and a few long bristles on the anterior part of the haltere (Figure 4.1H'). However, the severity of this morphological change looks much less pronounced than *miR-304*. It seems that *miR-10* overexpression has larger haltere (Figure 4.1L) with thick and long bristles on the anterior and distal part of capitellum (Figure 4.1L').

miR-92a, *miR-92b*, *miR-310*, *miR-311* and *miR-312* all belong to the same miRNA family (*miR-92* family), and they all induce classic haltere-to-wing transformation. Nevertheless, even though they have the same seed sequence

(Figure 4.2B), the morphological changes induced by the ectopic expression of miRNAs in this family vary (Figure 4.1). Misexpressions of *miR-92a* and *miR-92b* (Figure 4.1F, F', G, G') looks to have a similar phenotype, while ectopic expression of *miR-312* seems to have slightly bigger halteres with similarly placed ectopic bristles on the anterior of the haltere (Figure 4.1K, K') compared to those of *miR-92a* and *miR-92b*. It looks like that ectopic expression of *miR-310* in haltere discs causes a much more severe phenotype with bigger haltere size and several rows of ectopic bristles (Figure 4.1 I, I') compared to that of *miR-92a*, *miR-92b* and *miR-312*. Ectopic expression of *miR-311* seems to have the strongest haltere phenotypic change in this miRNA family, which can in some cases reach roughly three times the size of the control haltere, with an excess of ectopic bristles on the anterior part of the haltere (Figure 4.1 J, J'), even when compared to the ectopic bristles induced by *miR-iab-4* overexpression.

4.2.2 The molecular basis of miRNA effects on developing haltere

Following the phenotypical analysis of Class I miRNAs, it begs the question: what causes the different degree of homeotic transformations that I observed for different miRNAs? I suggest that different miRNAs may repress *Ubx* expression to different extents leading to a differential phenotypic severity in homeotic transformations.

One the one hand, *Ubx* effects on haltere morphology relies on the dosage of *Ubx*. Different *Ubx* mutant alleles lead to different *Ubx* expression levels in the developing haltere imaginal disc, being directly responsible for the varying severity of homeotic transformations in the adult appendage (Figure 1.4B) (Bender et al. 1983). *miR-iab-4*, *miR-303*, *miR-304*, *miR-252*, and *miR-10* may have different endogenous concentrations in the haltere (Figure 3.9), different target affinity to *Ubx* target sites and/or different number of target sites on the *Ubx* 3' UTR, which may lead to different levels of post-transcriptional repression. On the other hand, these miRNAs may regulate *Ubx* expression at different time points. It has been reported that modulating *Ubx* expression level at different time points during larval development leads to the differential specification of

particular cell-types of the adult appendages (Roch & Akam 2000). The moment at which *Ubx*-Gal4^{LDN} drives the expression of miRNAs could, for instance, enhance the ability of *miR-iab-4* to repress *Ubx* expression when compared to the remaining 4 miRNAs. Thus, I suggest that the differential repressing abilities of miRNAs on *Ubx* expression lead to the observed differential severities in homeotic transformations.

The other five miRNAs in Class I belong to the *miR-92* family. miRNAs that belong to the same miRNA family usually have a similar biological function because they have identical seed sequences and often share transcriptional regulation (when in a cluster), all of which suggest similar modulating roles of the same mRNA targets and therefore similar biological function. However, the morphological changes induced by the ectopic expression of miRNAs in the *miR-92* family vary as described in the last section (Figure 4.1). The mechanism underlying the observed phenotypic variations between miRNAs in the same family is unknown. I suggest that the seed of miRNA is not the only factor determing miRNA regulation and that other factors affect the miRNA regulation may be involved, which will be illustrated in the next section.

4.2.3 *miR-92* family

Although *miR-92a*, *miR-92b*, *miR-310*, *miR-311* and *miR-312* all belong to the *miR-92* family, these miRNAs induce different levels of homeotic transformations upon overexpression. As suggested in the last section, this points to the possibility that the seed of a given miRNA is not the only factor determing the outcome of miRNA regulation and consequently, that other factors may be involved in miRNA regulation. Here, I explore which factors may lead to differential miRNA regulatory effects within the same miRNA family. To understand the origin of variations in the phenotypic outcome among miRNAs in *miR-92* family, I analysed different features of miRNAs in the *miR-92* family from the perspectives of evolutionary conservation, mature miRNA sequences outside the seed and miRNA expression levels in white pre-pupal haltere imaginal discs (Figure 4.2).

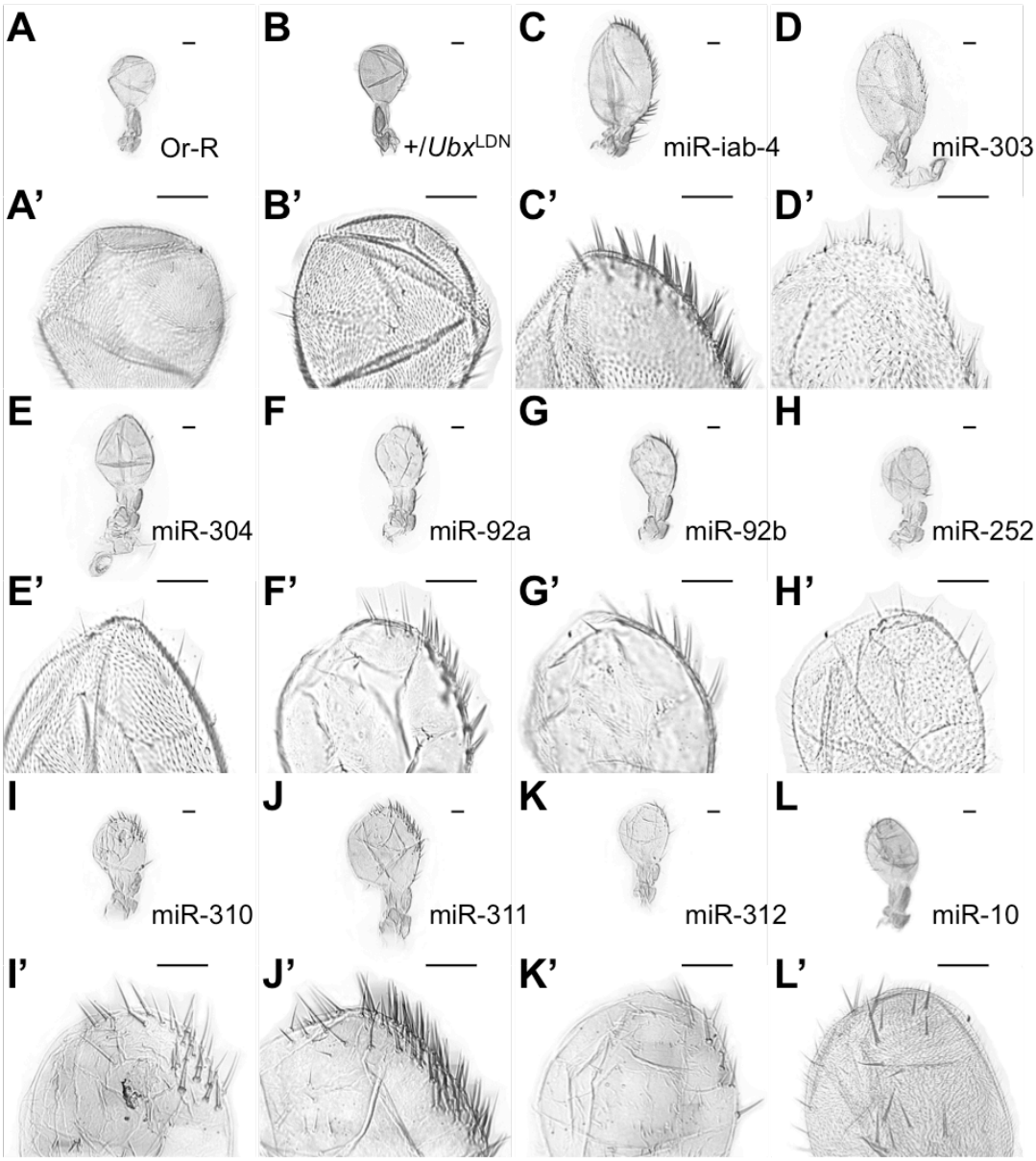


Figure 4.1 Summary of haltere phenotypes caused by ectopic expression of Class I miRNAs.

Images of haltere samples dissected from female flies of each genotype. (A-L) Haltere samples were observed under 10X magnification using a Leica DMRB microscope. (A'-L') Haltere samples (40X). (A, A') Or-R. (B, B') y w; +; +/*Ubx*>GAL4^{LDN}. (C, C') y w; +; UAS-miR-iab-4/*Ubx*>GAL4^{LDN}. (D, D') y w; +; UAS-miR-303/*Ubx*>GAL4^{LDN}. (E, E') y w; +; UAS-miR-304/*Ubx*>GAL4^{LDN}. (F, F') y w; +; UAS-miR-92a/*Ubx*>GAL4^{LDN}. (G, G') y w; +; UAS-miR-92b/*Ubx*>GAL4^{LDN}. (H, H') y w; +; UAS-miR-252/*Ubx*>GAL4^{LDN}. (I, I') y w; +; UAS-miR-310/*Ubx*>GAL4^{LDN}. (J, J') y w; +; UAS-miR-311/*Ubx*>GAL4^{LDN}. (K, K') y w; +; UAS-miR-312/*Ubx*>GAL4^{LDN}. (L, L') y w; +; UAS-miR-10/*Ubx*>GAL4^{LDN}. Class I miRNAs showing homeotic transformations, albeit with varying degrees. Scale bar represents 50μm.

4.2.3.1 Evolutionary conservation

Phylogenetic analysis of the *miR-92* family (Figure 4.2A) revealed some interesting features of this miRNA family. *miR-92a*, *miR-92b* and *miR-312*, which lead to milder morphological changes, are clustered together, while *miR-310* and *miR-311* which lead to more obvious *Ubx* mutant phenotypes are branched together. Thus, these results indicate that miRNAs that play similar roles in post-transcriptional regulation tend to cluster together.

4.2.3.2 Mature miRNA sequence

Figure 4.2B shows a multiple sequence alignment for miRNAs in the *miR-92* family. Interestingly, I find that *miR-310*, *miR-311* and *miR-312* share the same nucleotide 'U' (Uracil) in the g1 (position one of the mature miRNA sequence), while *miR-92a* and *miR-92b* have C (Cytosine) and A (adenosine) nucleotides, respectively. Schirle *et al.* reported that g1 inserts into a narrow pocket between the MID and L2 domains of Ago2 that specifically recognize A nucleotides (Schirle et al. 2014); as such the affinity between a target RNA with a t1 (position one of miRNA target site) A bound to Ago2 is almost three fold higher than equivalent targets with U, G (Guanine) or C nucleotides. Additionally, it has been shown that t1As are conserved in many mammalian miRNA target sites and enhance miRNA-mediated repression (Bartel 2009; Lewis et al. 2005). Thus, a highly conserved t1A in the mRNA target site with high miRNA target affinity may explain why g1U in the *miR-310*, *miR-311* and *miR-312* enhances post-transcriptional repression when compared to other members of the *miR-92* family (*miR-92a* and *miR-92b*), causing more obvious haltere-to-wing transformations.

Compared to other miRNAs in the *miR-92* family, *miR-311* and *miR-310* share a g9 (position nine of mature miRNA sequence) A and g11 (position eleven of mature miRNA sequence) U. As such, I wondered whether this feature in *miR-310* and *miR-311* might account for the more severe phenotypic changes observed, when compared to the ectopic expression of the rest three miRNAs (Figure 4.2B). However, as Schertel and colleagues found that nucleotides at

the position eleven of the mature miRNA do not significantly affect the target recognition of the *miR-92* family miRNAs (Schertel et al. 2012), therefore, it is unlikely that a g11U increases target affinity for miRNAs *miR-310* and *miR-311*.

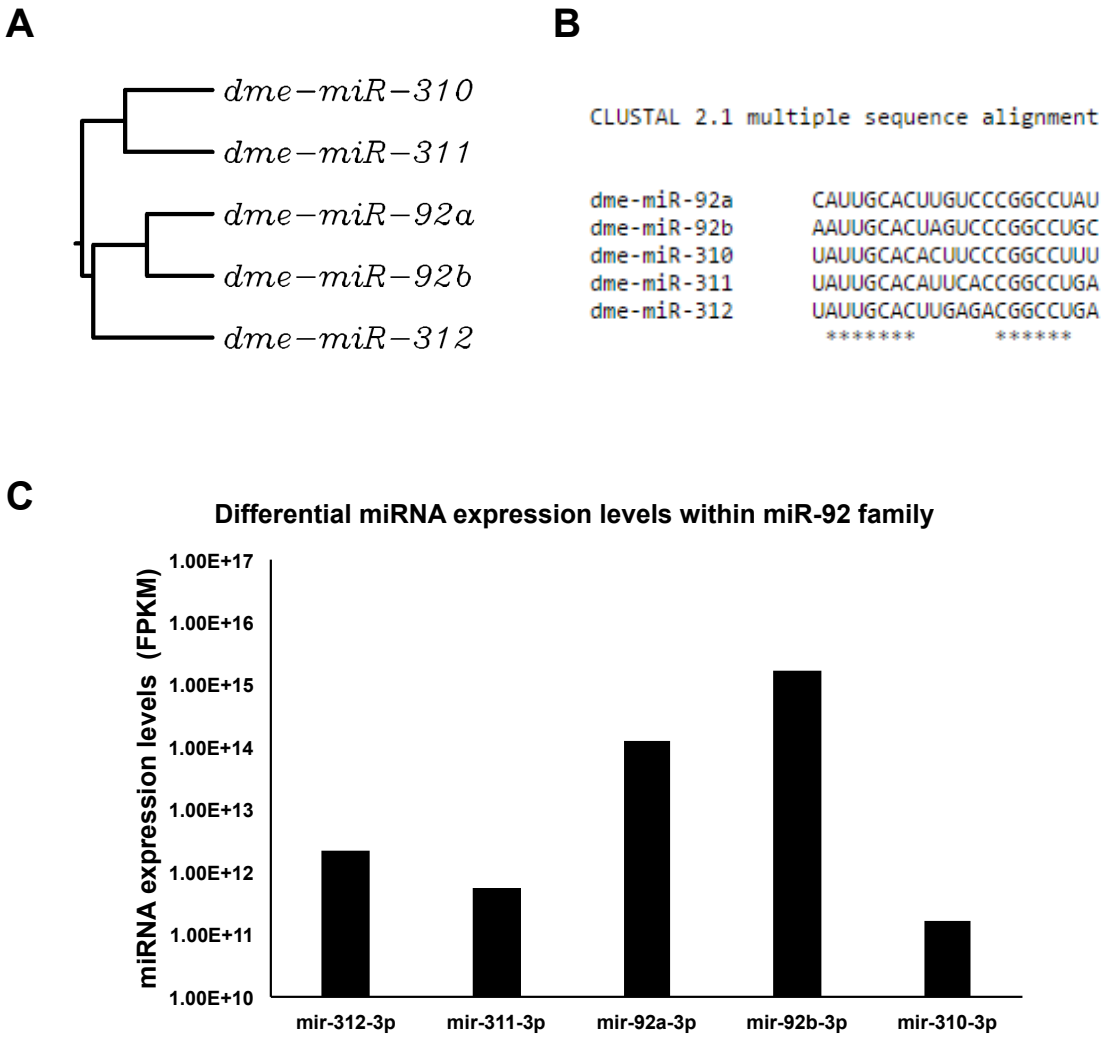


Figure 4.2 Analysis of the *miR-92* family.

(A) Rooted phylogenetic tree (UPGMA, with branch length) for the *miR-92* family. I found that *miR-310* and *miR-311*, which produce more severe haltere morphological changes, clustered together while the remaining three miRNAs cluster together in a separate branch, which implies that similar mature miRNA sequences lead to similar miRNA functions. (B) Alignment of mature miRNAs in the *miR-92* family using Cluster 2.1. All the mature miRNA sequences used here correspond to the 3p miRNA in each locus; '*' denotes conserved residues among all five mature miRNAs. (C) miRNA expression levels of the *miR-92* family in white pre-pupal haltere imaginal discs. The miRNA expression levels are calculated by FPKM. *miR-92a* and *miR-92b* have much higher endogenous expression levels in haltere imaginal discs, but their ectopic expression causes milder haltere morphological change compared to the rest three miRNAs in *miR-92* family. The *miR-313*, member of the *miR-92* family, has not been included in the analysis as the UAS stock for this miRNA was not available.

Lewis *et al.* reported that A nucleotides at t9 (position nine of the miRNA target site) can enhance the accuracy of miRNA targeting predictions by increasing the number of predicted miRNA-RNA relationships without a proportional increase in the number of false positives (Lewis et al. 2005). The g9U of *miR-92a*, *miR-92b* and *miR-312* could well complement with t9A. In this way, instead of causing more obvious morphological changes, *miR-310* and *miR-311* should have lower target affinity compared to *miR-92a*, *miR-92b* and *miR-312*, which is contrast with what I observed in our study.

Here, I suggest two possibilities that may explain this paradox. One possibility is that the affinity enhancement by having A nucleotides at t1 is more important compared to A nucleotides at t9. Thus, the target affinity "boost" due to a 'U' in the g1 position of *miR-310* and *miR-311* sequence overrides the enhancement by the occurrence of a 'U' in the g9 position of *miR-92a* and *miR-92b* sequences, causing more obvious haltere-to-wing transformations. Another possibility is that good complementarity between miRNA and mRNA at g9 may instead decrease target affinity. Schirle et al. reported that sequence complementary in g8 (position eight of miRNA sequence) enhances the affinity of Ago2 to target RNAs (Schirle et al. 2014). However, the extended complementary in g9 and g10 (position ten of mature miRNA sequence) cannot further increase affinity, and g9 pairing can even be detrimental to the stability of the Ago complex, as the pairing beyond g9 of miRNA sequence needs to further open the Ago2 central cleft, and such conformational changes decrease target affinity.

To test my hypothesis, I suggest the following experiments. Single nucleotide mutations can be introduced by CRISPR to make a mutant version of the miRNA and observe how a single nucleotide mutation affects miRNA target recognition. For example, I can replace the U nucleotides at g1 of *miR-310*, *miR-311* and *miR-312* for an A or C, or change the position one nucleotides of *miR-92a* and *miR-92b* to U, in order to determine how these mutations affect the observed haltere morphological changes. If the severity of haltere morphological phenotype is reduced by a g1U change to an A or C in *miR-310*, *miR-311* and *miR-312*, I will conclude that a g1U can increase the affinity of

miRNA-*Ubx* interactions. Conversely, if the mutation of nucleotides in position one of *miR-92a* and *miR-92b* leads to an increasing of the severity of homeotic transformations, I can also confirm the importance of g1U for miRNA-*Ubx* target affinity. Similarly, manipulations of the nucleotide identity of g9 and g11 can be attempted to test whether the significance of pairing in these two positions affects target affinity.

4.2.3.3 miRNA expression levels

The endogenous miRNA expression level contributes to miRNA target affinity. It has been suggested that in a low miRNA abundance condition, miRNAs may only bind to high affinity target sites, while high miRNA abundance may saturate all the high affinity target sites, allowing free miRNA molecules to bind moderate affinity target sites, thus enhancing target repression (Coronnello et al. 2012). Thus, miRNAs with higher expression levels should cause more severe repression when compared to those in the same miRNA family with relatively lower expression. For example, *miR-310*, *miR-311* and *miR-312* cause much more severe homeotic transformation compared to *miR-92a* and *miR-92b*, thus their expression level should be higher. However, when I analysed miRNA expression levels in the *miR-92* family (Figure 4.2A), I found that the endogenous expression levels of *miR-92a* and *miR-92b*, which cause milder phenotypic changes, are much higher than those of the remaining three miRNAs. One possibility for this disagreement is that the number of miRNA target site for each miRNA varies among miRNAs in the *miR-92* family. *miR-310*, *miR-311* and *miR-312* may have more target sites than *miR-92a* and *miR-92b* and ectopic expression of these miRNAs may provide sufficient levels to allow them to bind to all target sites; thus, *miR-310*, *miR-311* and *miR-312* would have much stronger phenotypic changes compared to *miR-92a* and *miR-92b*. Another possibility is the target affinity is different among miRNAs in *miR-92* family. If the target affinity for *miR-310*, *miR-311* and *miR-312* is much higher than that of *miR-92a* and *miR-92b*, the ectopic expression of *miR-310*, *miR-311* and *miR-312* is expected to trigger much more severe homeotic transformations.

4.2.4 Phenotypic analysis of miRNAs from Class I by scanning electronic microscope (SEM).

Due to the easy and cheap preparation procedure and the possibility of being imaged under high magnification microscope, *Drosophila* haltere cuticle preparation is a very important technique for the overview of the haltere morphological change. However, the flat preparation of the haltere cuticle loses the information of the three-dimension (3-D) structure of the haltere, and makes it impossible to observe the haltere surface topography and composition. But, with the help of the Scanning Electron Microscopy (SEM) technique this problem can be solved. I observed haltere samples from phenotypic Class I using SEM to further analyse the haltere morphological changes caused by ectopic expression of miRNAs in higher detail (Figures. 4.3- 4.5).

It looks like that the haltere of *miR-iab-4* ectopic expression flies (Figure 4.3B) is almost six times the size of the control (Figure 4.3A). Even though the *miR-iab-4* induced haltere change resembles a small wing, the organ is not as flat as the wing. Consequently, there is still low expression of Ubx protein after ectopic expression of *miR-iab-4* (Figure 4.8B'). Otherwise, complete deletion of *Ubx* leads to two pairs of wings (Lewis 1978). Furthermore, on the anterior part of the haltere (Figure 4.3B'), I can observe at least duplicated rows of the wing-like margin bristles. Also, the proximal part of the haltere (Figure 4.3C, C') has a flattened wing shape with wing margin bristles. As the size of the haltere increases, the small trichomes covering the haltere (Figure 4.3B'') also increase in size, when compared to the control (Figure 4.3A').

The haltere induced by the overexpression of *miR-303* seems to have the shape of a flat small wing (Figure 4.3D) with duplicated rows of wing margin bristles on the anterior part of the haltere (Figure 4.3D'). More interestingly, this haltere shows transformations in cell type due to the ectopic expression of *miR-303* (Figure 4.3D''), changing from haltere-like cuboidal cells into wing-like star cells (Roch et al. 2003; Pavlopoulos & Akam 2011). The small bristles covering the haltere surface are similar to those of the wing (Roch & Akam 2000). When

this phenotypic information is compared to the well-known evidence of *Ubx* regulation by *miR-iab-4* (Ronshaugen et al. 2005; Stark et al. 2008), it suggests that *miR-303* could be a good candidate for the regulation of *Ubx* expression.

It seems that the haltere phenotypes induced by ectopic expression of *miR-252* are quite similar to those resulting from *miR-10* overexpression. They both have reversed trichomes (Figure 4.4A''-B'') covering the surface of the haltere and one row of ectopic bristles (Figure 4.4A'-B') on the anterior of the haltere. However, the three-dimensional shape of the halteres induced by these two miRNAs is different. Haltere shapes induced by the ectopic expression of *miR-252* approximate a half cut haltere with the size looks bigger than the control (Figure 4.3A, 4.4A), while those induced by misexpression of *miR-10* seems to have a similar shape to the control (Figure 4.3A, 4.4B), albeit slightly bigger. It looks like that the *miR-304* overexpression haltere is bigger (Figure 4.4C) than *miR-252* (Figure 4.4A) and *miR-10* (Figure 4.4B), with one row of ectopic bristles on the anterior part of the haltere but no reversed trichomes. It seems that this row of ectopic bristles and the size of the hairs covering the haltere all increase with haltere size (Figure 4.4C') when compared to those of *miR-252* (Figure 4.4A') or *miR-10* (Figure 4.4B').

It is interesting to note that ectopic expression of *miR-92a*, *miR-92b*, *miR-311*, *miR-312*, all members of the *miR-92* family (see previous sections), leads to similar phenotypes showing fewer small bristles covering the haltere surface (Figure 4.5A''-D''). This is consistent with previous reports showing that ectopic expression of *miR-92a* and its seed relatives have the unique ability to trigger trichome loss (Bejarano et al. 2012; Schertel et al. 2012). Haltere and wing both have trichomes covering the surface of these two organs, so this kind of phenotypic change is probably involved in the regulation of genes other than *Ubx*. As all these miRNAs share the same seed sequence, it is thus likely that they have the ability to target the same mRNAs. It has been shown that *miR-92a* controls "naked valley" size in the legs by repressing *shavenoid* (Arif et al. 2013). Thus, I suggest that the trichome loss on the haltere by ectopic expression of miRNAs from *miR-92* family could occur by downregulation of *shavenoid*, similarly to what has been reported for the leg (Arif et al. 2013).

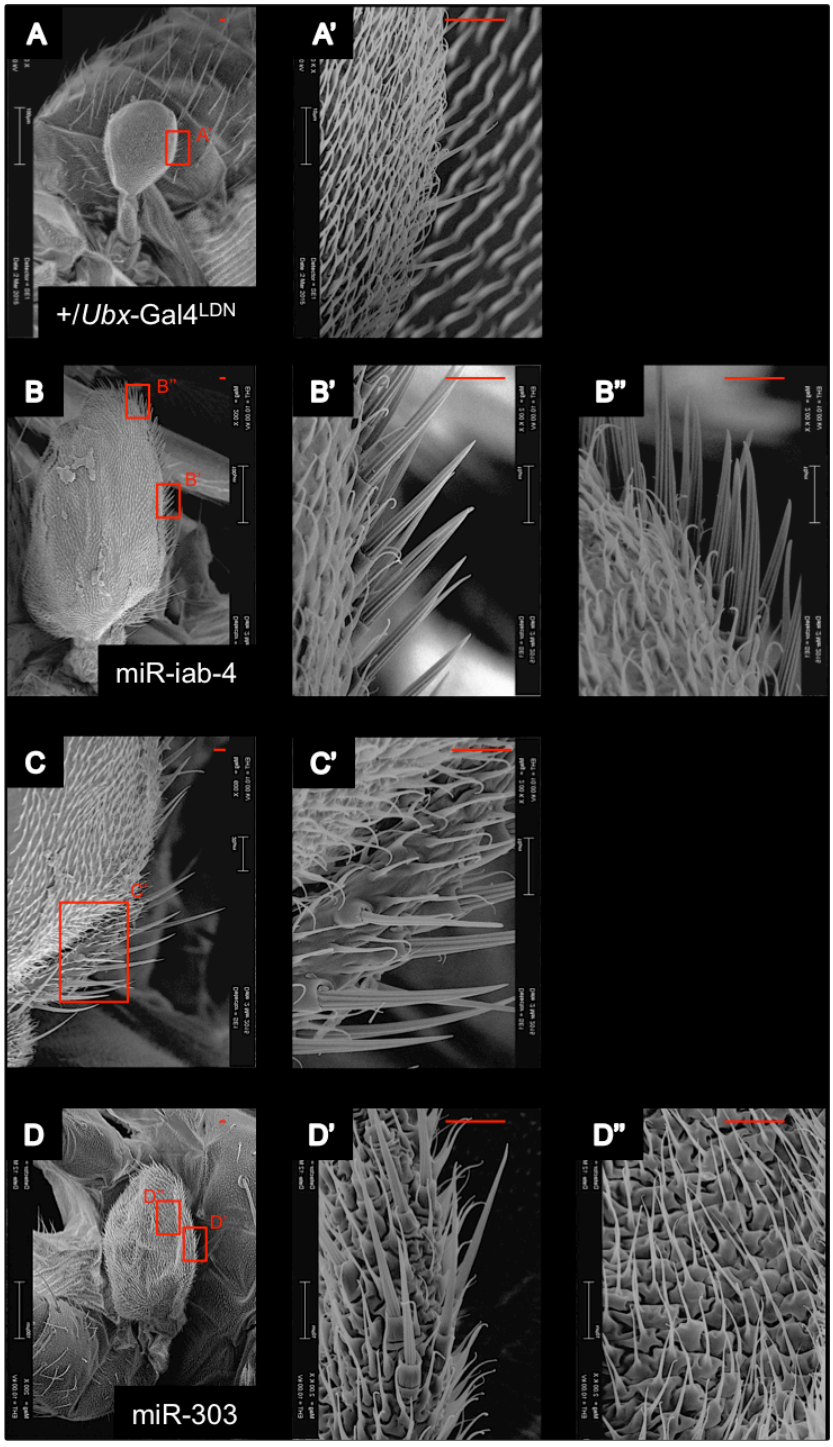


Figure 4.3 Analysis of haltere phenotypes caused by overexpression of miRNAs from Class I using SEM (Part I).

Images of haltere samples are from female flies of each genotype. (A, B, D) Haltere samples were taken under 200 times magnification by using Leo stereo scan 420 scanning electron microscope. (A'-D', B'', D'') Haltere samples under 2000 times magnification. (C) Haltere sample of y w; +; UAS-miR-iab-4/*Ubx*>GAL4^{LDN} scanned under 600 times magnification. (A, A') y w; +; +/*Ubx*>GAL4^{LDN}. (B-B'', C, C') y w; +; UAS-miR-iab-4/*Ubx*>GAL4^{LDN}. (D-D'') y w; +; UAS-miR-303/*Ubx*>GAL4^{LDN}. Red rectangles denote areas that are imaged with higher magnification, and the corresponding images are shown on the right of the images. Scale bar represents 10μm.

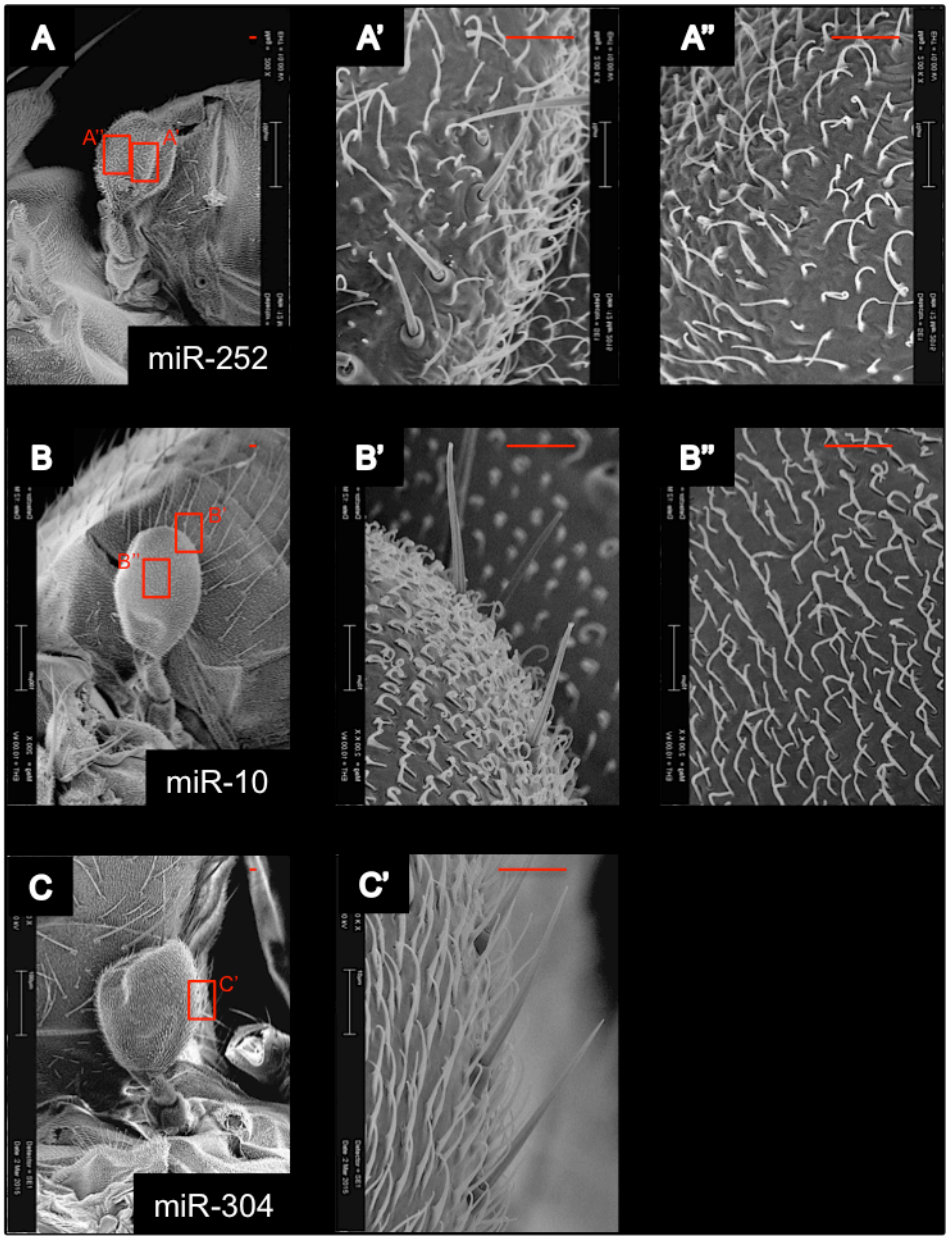


Figure 4.4 Analysis of haltere phenotypes caused by overexpression of miRNAs from Class I using SEM (Part II).

Images of haltere samples are from female flies of each genotype. (A-C) Haltere samples were taken under 200 times magnification by using Leo stereo scan 420 scanning electron microscope. (A'-C', A''-B'') Haltere samples under 2000 times magnification. (A-A'') y w; +; UAS-miR-252/*Ubx*>GAL4^{LDN}. (B-B'') y w; +; UAS-miR-10/*Ubx*>GAL4^{LDN}. (C, C') y w; +; UAS-miR-304/*Ubx*>GAL4^{LDN}. Red rectangles denote areas that are imaged with higher magnification, and the corresponding images are shown on the right of the images. Scale bar represents 10μm.

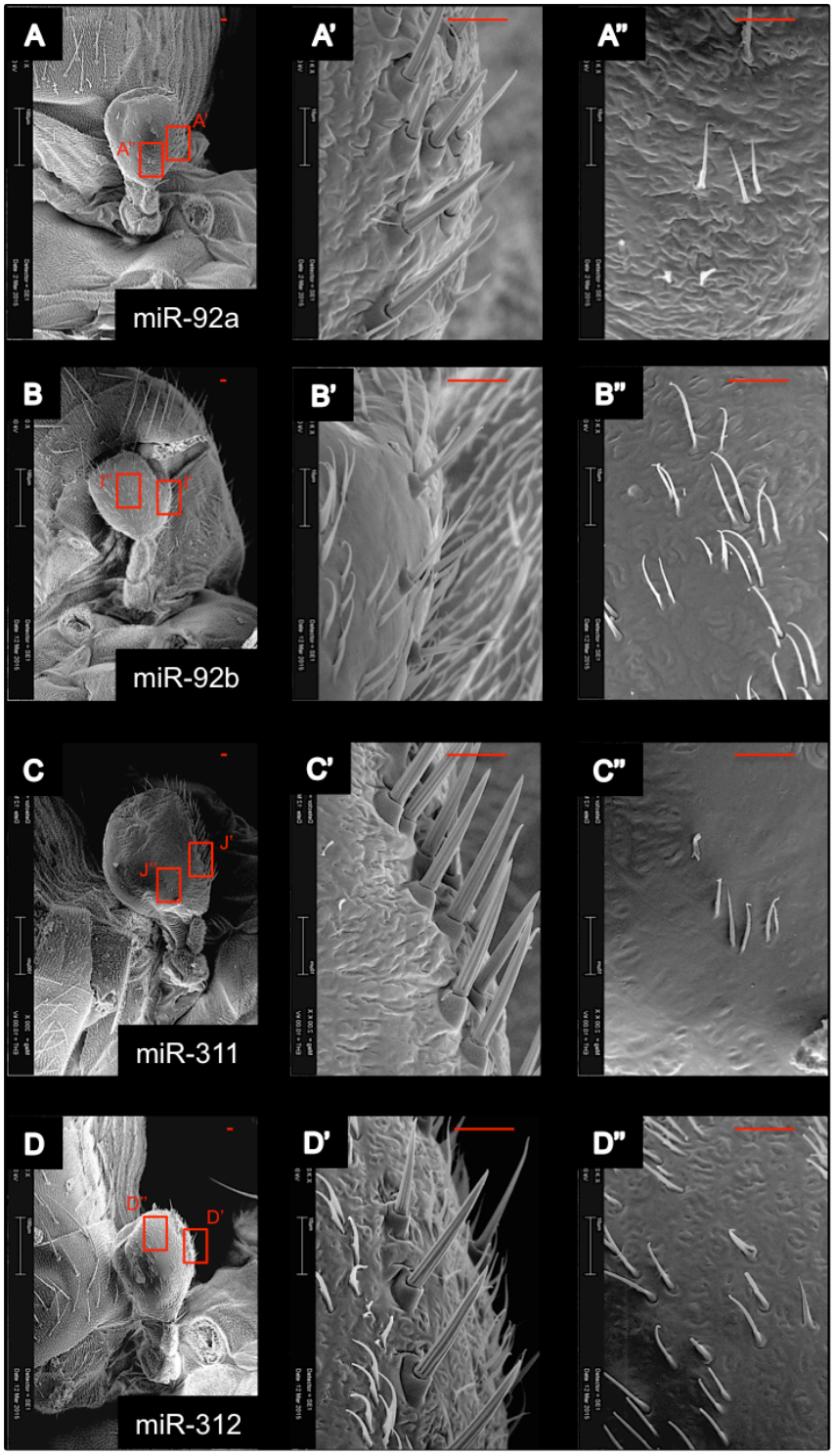


Figure 4.5 Analysis of haltere phenotypes caused by overexpression of miRNAs from Class I using SEM (Part III).

Images of haltere samples are from female flies of each genotype. (A-D) Haltere samples were taken under 200 times magnification by using Leo stereo scan 420 scanning electron microscope. (A'-D', A''-D'') Haltere samples under 2000 times magnification. (A-A'') y w; +; UAS-miR-92a/*Ubx*>GAL4^{LDN}. (B-B'') y w; +; UAS-miR-92b/*Ubx*>GAL4^{LDN}. (C-C'') y w; +; UAS-miR-311/*Ubx*>GAL4^{LDN}. (D-D'') y w; +; UAS-miR-312/*Ubx*>GAL4^{LDN}. Red rectangles denote areas that are imaged with higher magnification, and the corresponding images are shown on the right of the images. Scale bar represents 10μm.

The haltere shapes induced by *miR-92a*, *miR-92b* and *miR-312* (Figure 4.5A, B, D) seem similar to that of the control (Figure 4.3A), albeit with bigger volumes. However, the haltere shape induced by *miR-311* (Figure 4.5C) looks similar to the shape of wing and much bigger than the control (Figure 4.3A). The haltere size changes due to the ectopic expression of miRNAs in the *miR-92* family can be summarized as *miR-311*> *miR-312*> *miR-92a*>=*miR-92b* (Figure 4.5A-D), consistent with what I observed in haltere cuticle preparations (Figure 4.1 F-G, J-K). The size of ectopic bristles induced by overexpression of *miR-92a*, *miR-92b*, *miR-311* and *miR-312* is well correlated with the size changes in the adult halteres (Figure 4.5 A'-D'). However, overexpression of *miR-92a* and *miR-311* induces three rows of wing margin bristles (Figure 4.5 A', C'), while ectopic expression of *miR-92b* and *miR-312* develops one row and two rows of wing margin bristles, respectively (Figure 4.5 B', D').

4.2.5 miRNAs from phenotypic Class I gain-of-function result in phenotypic changes linked to *Ubx* loss-of-function in the third instar larval haltere imaginal discs.

In the previous section, I show that gain-of-function experiments for some miRNAs result in haltere-to-wing transformations (phenotypic Class I). However, are these miRNAs inducing haltere morphological changes through the direct regulation of *Ubx* expression? If yes, to what extent do these miRNAs affect *Ubx* levels thus causing these changes? To answer the questions above, I performed antibody stainings for Ubx in haltere imaginal discs from third instar larvae, and I quantified Ubx expression level changes in different Class I miRNA genotypes (Figure 4.7-4.9). As Ubx is mainly expressed in the haltere pouch (Figure 4.6B), an aspect that is recapitulated by the driver used in these experiments, *Ubx*-Gal4^{LDN} (Figure 4.6D), I quantified Ubx expression level changes in the haltere pouch in order to make the result more easily interpretable. The yellow circle in Figure 4.6A illustrates the area in the haltere imaginal disc that was used for the quantifications.

With the exception of *miR-10* (Figure 4.7D', F), all miRNAs from phenotypic

Class I repress Ubx expression in the haltere imaginal discs (Figure 4.7B', C', E', F, 4.8B'-G', 4.9). The average expression level of Ubx is decreased by ectopic expression of *miR-10* (Figure 4.7D') when compared to control imaginal discs (Figure 4.7A'), but no significant difference is detected (Figure 4.7F).

The reduction of Ubx expression level by ectopic expression of *miR-iab-4* is the most dramatic (Figure 4.8B', 4.9), a result which is consistent with the haltere morphological changes shown in Figure 4.4B. Likewise, I found that ectopic expression of miRNAs *miR-311*, *miR-312* and *miR-92b* (Figure 4.8E'-G', 4.9) all lead to a significant decrease in Ubx expression. In the case of *miR-252* (Figure 4.4A) that shows less severe phenotypic changes when compared to *miR-iab-4*, *miR-92b*, *miR-311* and *miR-312* (Figure 4.3B, 4.5B-D), I found that the observed decrease in Ubx levels (Figure 4.8C', 4.9) is less pronounced than in all remaining cases. Thus, I conclude that the level of Ubx repression by miRNAs appears to be positively correlated with the severity of haltere homeotic transformations in most cases.

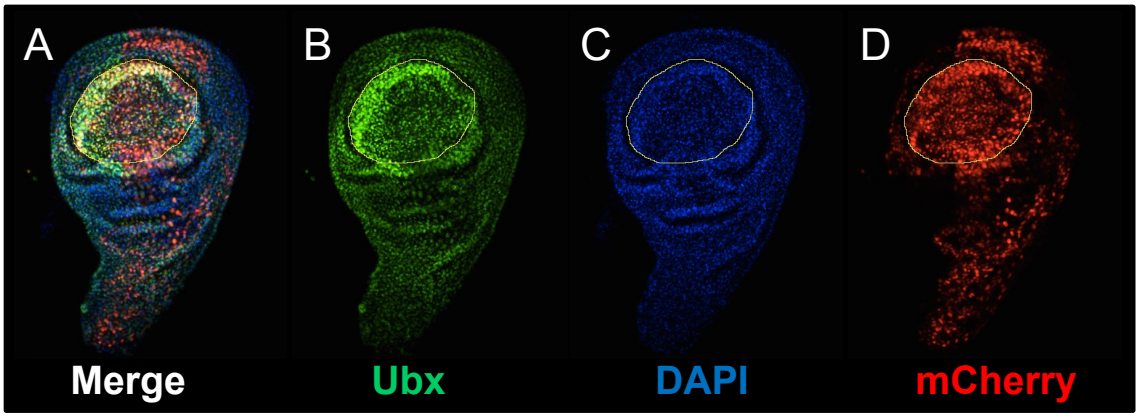


Figure 4.6 Identification of a subregion of the haltere imaginal disc for expression analysis.

Immunohistochemistry stainings in *Ubx*-Gal4^{LDN}.UAS-mCherry third instar larval haltere imaginal disc. (A) Merged image for B, C and D. (B) Antibody staining for Ubx. (C) DAPI staining, denoting individual cell nuclei. (D) Antibody staining for mCherry. The circled area in the haltere pouch is the area used for quantification in Figure 4.7, 4.8, 4.9. Ubx is mainly expressed in the haltere pouch and colocalized with the *Ubx*-Gal4^{LDN} expressing cells in this area. I quantified Ubx expression changes in the haltere pouch in order to assess the miRNA effects on Ubx expression.

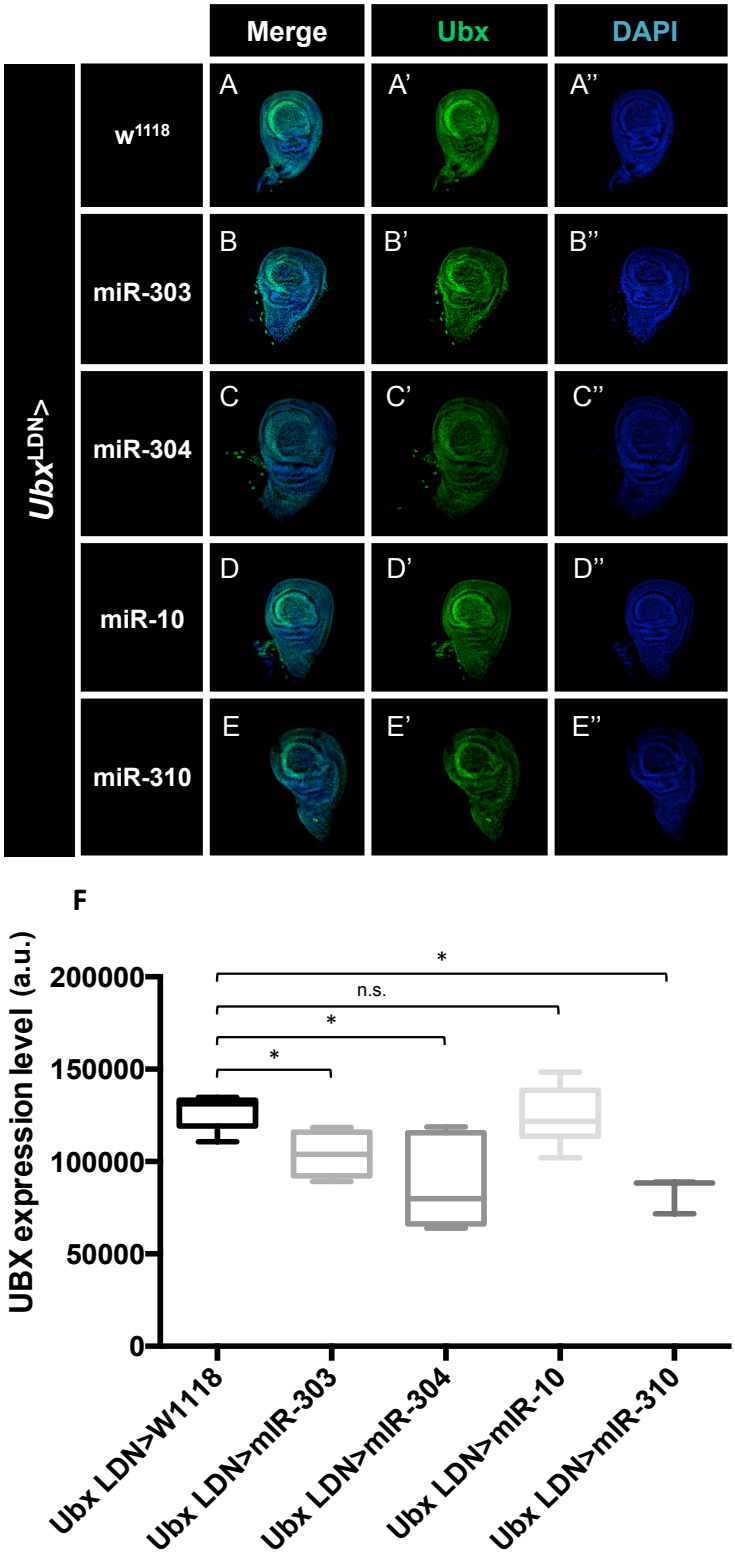


Figure 4.7 Immunohistochemistry analysis of Ubx expression levels in haltere imaginal discs after ectopic expression of miRNAs from phenotypic Class I (Part I).

(Legend on the following page)

Figure 4.7 Immunohistochemistry analysis of Ubx expression levels in haltere imaginal discs after ectopic expression of miRNAs from phenotypic Class I (Part I). Analysis of Ubx protein expression in third instar larval haltere imaginal discs overexpressing Class I miRNAs. (A-E) Merged images for Ubx and DAPI staining. (A'-E') Antibody staining for Ubx. (A''-E'') DAPI staining. (A-A'') y w1118; +/UAS-mCherry; +/*Ubx*>GAL4^{LDN}. (B-B'') y w; +/UAS-mCherry; UAS-miR-303/*Ubx*>GAL4^{LDN}. (C-C'') y w; +/UAS-mCherry; UAS-miR-304/*Ubx*>GAL4^{LDN}. (D-D'') y w; +/UAS-mCherry; UAS-miR-10/*Ubx*>GAL4^{LDN}. (E-E'') y w; +/UAS-mCherry; UAS-miR-310/*Ubx*>GAL4^{LDN}. The Ubx expression level quantifications were carried out using the method as described in Figure 4.6. (F) Ubx immunohistochemistry analysis by ectopic expression of *miR-303*, *miR-304*, *miR-10* and *miR-310*. The x-axis indicates the genotype of the sample. The *y*-axis denotes Ubx expression calculated by arbitrary fluorescence intensity units (a.u.). An average of 5 haltere imaginal discs of each genetype was used in all analysis. The staining intensities were quantified by ImageJ, and statistical analysis were performed using a Mann-Whitney test (unpaired *t*-test); P>0.05 (non-significant; n.s.), P<0.05 (*). This set of experiment only has one biological repeat. Ubx expression levels are significantly reduced after ectopic expression of any of the following 3 miRNAs in Class I: *miR-303*, *miR-304*, *miR-310*. This experiment is an independent experimental set from Figure 4.8.

However, the Ubx expression level reduction by ectopic expression of *miR-303* (Figure 4.7B', F) does not follow this rule. Even though it seems that the overexpression of *miR-303* (Figure 4.1F, 4.3D) in haltere imaginal discs leads more severe phenotypic change compared to *miR-304* and *miR-310* (Figure 4.1E,I, 4.4C), the Ubx expression level reduction looks far less than that of *miR-304* and *miR-310* ectopic expression (Figure 4.7C', E', F). One possible explanation is that in addition to regulating *Ubx* expression in the haltere imaginal discs, *miR-303* may also regulate *Ubx* targeted genes that control haltere morphology. For example, the formation of margin bristles is regulated by the proneural *achaete* (*ac*) and *scute* (*sc*) target genes and the haltere size and cell morphology are controlled by the *Drosophila Serum Response Factor* (*DSRF* or *blistered*), *vg* quadrant enhancer and *sal* target genes (Weatherbee et al. 1998). Thus, even though *miR-303* represses *Ubx* expression in a less efficient manner, *miR-303* could also regulate the aforementioned *Ubx* target genes, leading to an additive effect on morphological changes and thus similar functional result as miRNAs solely targeting *Ubx* with higher efficiency. However, the experimental overexpression of *miR-303* in haltere imaginal discs has only been performed once. Therefore, this experiment should be repeated in order to consolidate our observation.

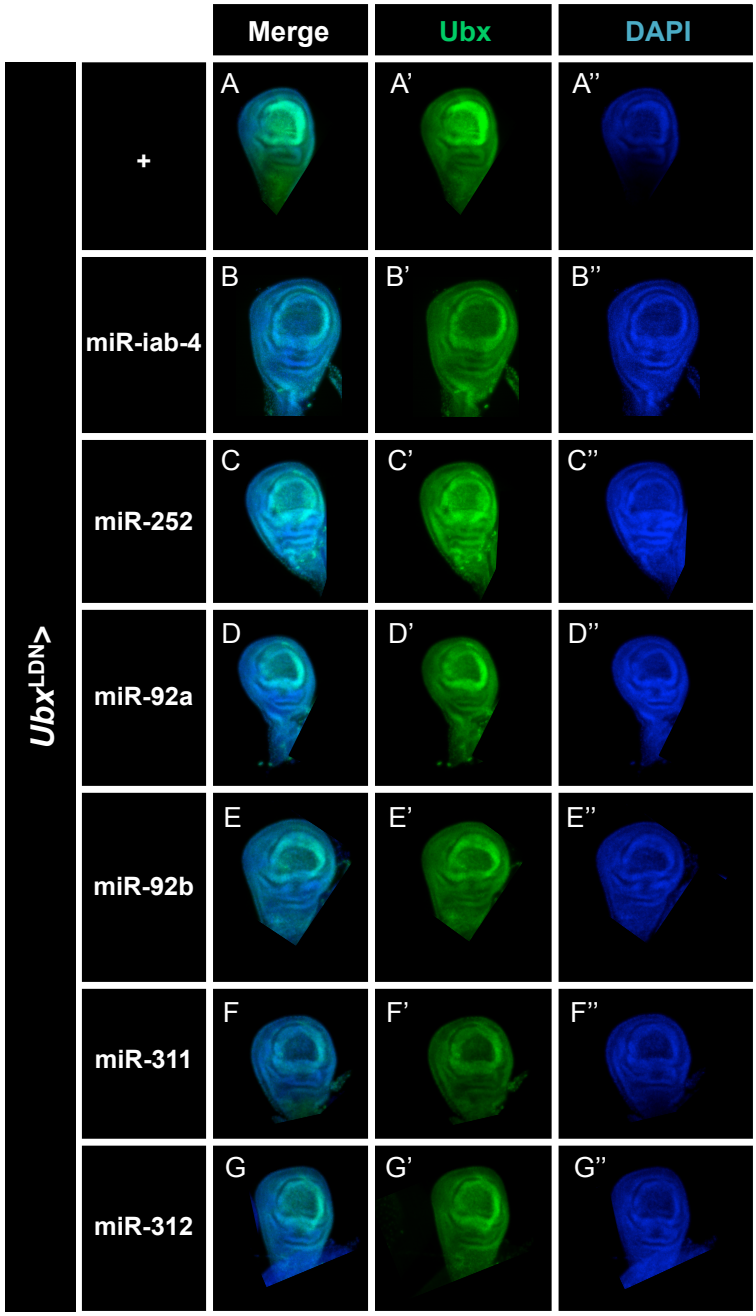


Figure 4.8 Immunohistochemistry analysis of Ubx expression levels in haltere imaginal discs after ectopic expression of miRNAs from phenotypic Class I (Part II).

Analysis of Ubx protein expression in third instar larval haltere imaginal discs overexpressing Class I miRNAs. (A-G) Merged images for Ubx and DAPI staining. (A'-G'') Antibody staining for Ubx. (A''-G'') DAPI staining. (A-A'') y w; +/UAS-mCherry; +/*Ubx*>GAL4^{LDN}. (B-B'') y w; +/UAS-mCherry; UAS-miR-iab-4/*Ubx*>GAL4^{LDN}. (C-C'') y w; +/UAS-mCherry; UAS-miR-252/*Ubx*>GAL4^{LDN}. (D-D'') y w; +/UAS-mCherry; UAS-miR-92a/*Ubx*>GAL4^{LDN}. (E-E'') y w; +/UAS-mCherry; UAS-miR-92b/*Ubx*>GAL4^{LDN}. (F-F'') y w; +/UAS-mCherry; UAS-miR-311/*Ubx*>GAL4^{LDN}. (G-G'') y w; +/UAS-mCherry; UAS-miR-312/*Ubx*>GAL4^{LDN}. This set of experiment has three biological repeats. The quantification of Ubx protein levels is shown in Figure 4.9. This experiment is an independent experimental set from Figure 4.7.

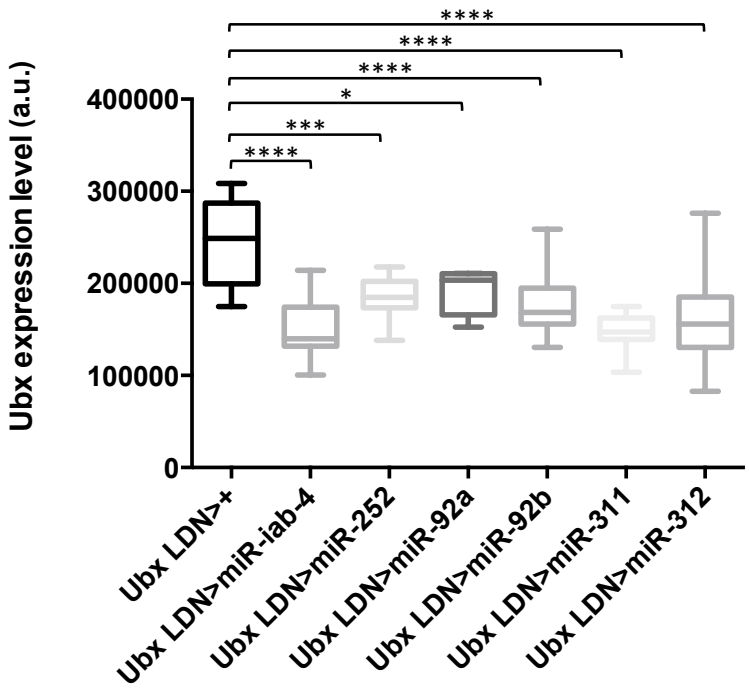


Figure 4.9 Immunohistochemistry analysis of Ubx expression levels in haltere imaginal discs after ectopic expression of miRNAs from phenotypic Class I (Part III).

Ubx expression level quantifications for Figure 4.8 were carried out using the method as described in Figure 4.6. Immunohistochemistry analysis of Ubx levels after ectopic expression of miRNAs, *miR-iab-4*, *miR-252*, *miR-92a*, *miR-92b*, *miR-311* and *miR-312*. The x-axis indicates the genotype of the samples. The *y*-axis denotes Ubx expression. An average of 12 haltere imaginal discs of each genetype was used in all analysis. The intensities of staining were quantified by ImageJ, and statistical analysis were performed using a Mann-Whitney unpaired *t*-test; P<0.05 (*), P<0.001 (***), P<0.0001 (****). A significant reduction in Ubx expression levels is observed after ectopic expression of any of the following six miRNAs in Class I: *miR-iab-4*, *miR-252*, *miR-92a*, *miR-92b*, *miR-311* and *miR-312*. In total, nine miRNAs are sufficient to repress Ubx expression in third instar larval haltere imaginal discs (Figure 4.7, 4.8, 4.9).

4.2.6 Validation of miRNAs that regulate Ubx expression by directly targeting *Ubx*.

In the previous section I showed that, using immunohistochemistry techniques, nine miRNAs from phenotypic Class I have the ability to regulate the expression of *Ubx* in haltere imaginal discs. However, it is fully possible that these miRNAs regulate *Ubx* indirectly i.e. by targeting upstream *Ubx*-activating genes, thus leading to the indirect repression of *Ubx*. Therefore, it is important to identify miRNAs that directly target *Ubx*, leading to Ubx expression changes that cause haltere-to-wing transformations.

To address this question, I used a specific *Ubx* 3' UTR transgene. As the *Ubx* long 3' UTR is the dominant *Ubx* 3' UTR isoform in the third instar larval haltere imaginal discs (Dr Richard Kaschula PhD thesis, Figure 3.1), I applied UAS-GAL4 strategy using a UAS-*Ubx* long 3' UTR transgenic stock to uncover miRNAs that directly target *Ubx in vivo*. UAS-*Ubx* long 3' UTR and UAS-miRNA were co-expressed in the haltere imaginal discs using the *sd*-GAL4 genetic driver (Figure 4.10B). In addition to this, I used a UAS-*K10* 3' UTR construct (*K10*, *female sterile (1) K10*) as a negative control with which to rule out of the possibility that our strategy does not lead to the repression of any 3' UTR (Figure 4.10A).

In addition to all nine miRNAs that regulate Ubx expression when ectopically expressed in third instar larval haltere imaginal discs, I have also studied *miR-9a* and *miR-310C*, a negative control and positive control, respectively. Ectopic expression of *miR-9a* shows no morphological changes (Figure 3.7RR), while *miR-310C* is sufficient and necessary to regulate *Ubx* expression in haltere imaginal discs (Richard Kaschula&Claudio Alonso, unpublished data). As expected, *miR-310C* directly targets *Ubx* long 3' UTR (Figure 4.15H'), while *miR-9a* does not target this sequence at all (Figure 4.15C').

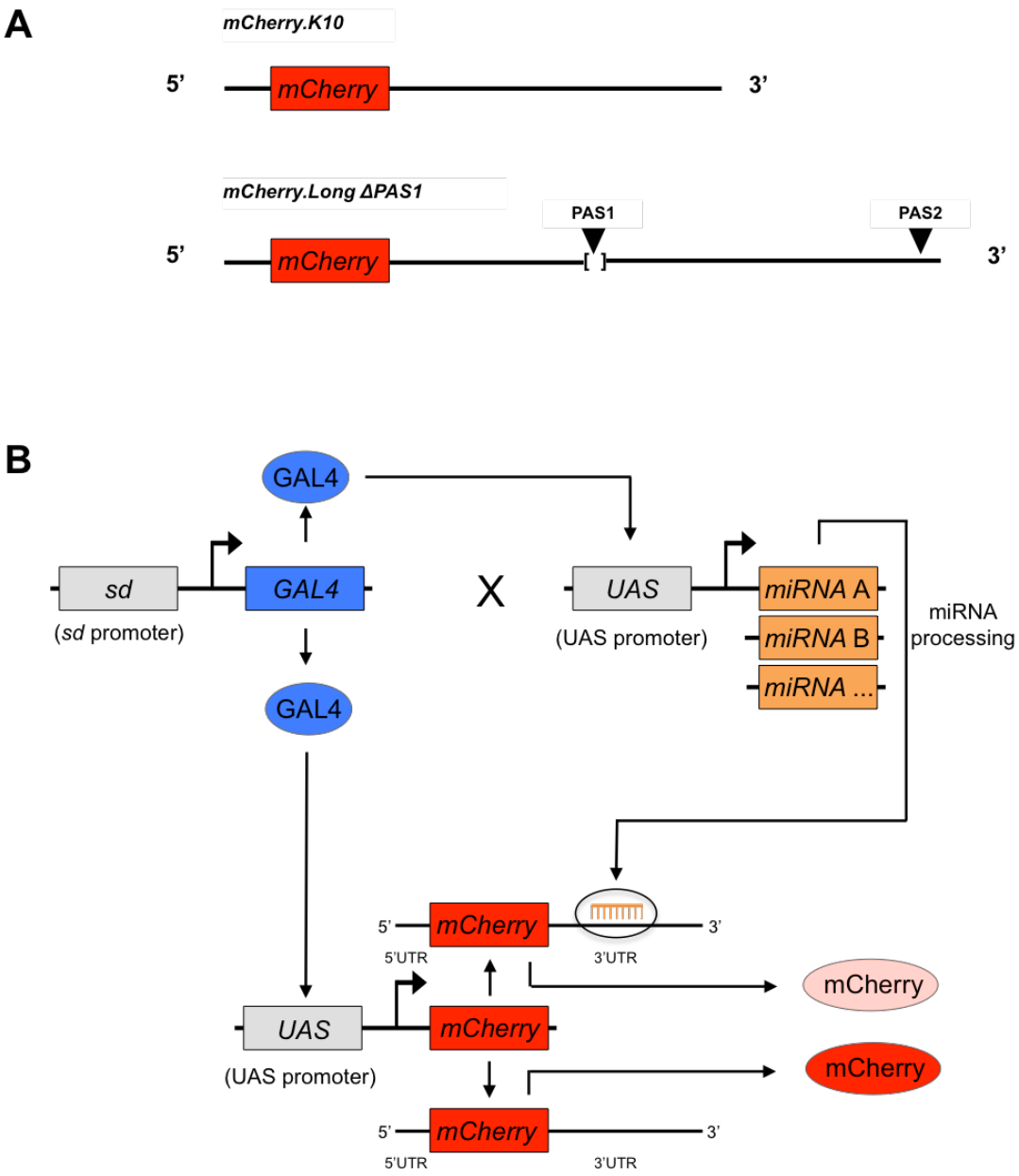


Figure 4.10 Use of 3' UTR fluorescent reporters to identify miRNA-mRNA interactions *in vivo*.

(A) Diagram depicting UAS-*mCherry K10* 3' UTR (used as control) and UAS-*mCherry Ubx* long 3' UTR (with PAS1 site deleted) constructs. (B) Diagram denoting genetic strategy to measure the expressions of mCherry in the haltere imaginal discs of the y w; UAS-*mCherry Ubx* long 3' UTR/+; UAS-miRNA/*sd*>GAL4 (genetic driver expressed in haltere imaginal discs) and y w; UAS-*mCherry K10* 3' UTR/+; UAS-miRNA/*sd*>GAL4 (Bejarano et al. 2012; Thomsen et al. 2010). If a direct interaction between miRNA and 3' UTR occurs, then the expression of mCherry will decrease compared to control.

I classified the miRNAs tested into three categories. The first category comprises miRNAs that directly target *Ubx* long 3' UTRs, but not *K10* 3' UTRs. For example, *miR-303*, *miR-312* and *miR-92a* appeared to directly regulate *Ubx* expression in third instar larval haltere imaginal discs (Figure 4.11B', C', D, 4.15D', I), while no discernable interactions with *K10* 3' UTRs were observed (Figure 4.12B', C', D, 4.16D', H). The second category consists of miRNAs that directly target the 3' UTR of both genetic constructs. For instance, *miR-310*, *miR-311*, *miR-iab-4* and *miR-310C* all directly target both *Ubx* long 3' UTRs and *K10* 3' UTRs (Figure 4.13B'-C', D, 4.15B', H', I, 4.14B'-C', D, 4.16B', G', H). Finally, I established a third category consisting of miRNAs that do not directly target *Ubx* long 3' UTRs at all, such as *miR-9a*, *miR-92b*, *miR-304* and *miR-252* (Figure 4.15C', E'-G', I). In summary, *miR-303*, *miR-310*, *miR-311*, *miR-312*, *miR-92a* and *miR-iab-4* have the ability to directly regulate *Ubx* expression in the *Drosophila* haltere.

However, the aforementioned results do not mean that these molecular interactions occur *in vivo*. In other words, if a given miRNA is indeed functional during haltere development, a necessary prerequisite is that this miRNA should be endogenously present in the haltere imaginal disc. In order to establish which of the studied miRNAs directly regulate *Ubx* mRNAs during haltere development, I measured the expression level of the six miRNAs that directly target *Ubx* 3' UTR. I found that miRNAs *miR-310-3p*, *miR-311-3p*, *miR-312-3p*, *miR-92a-3p*, and *miR-iab-4-5p* are present in the haltere (Figure 4.17). Thus, I suggest that these five miRNAs may play roles during the haltere development.

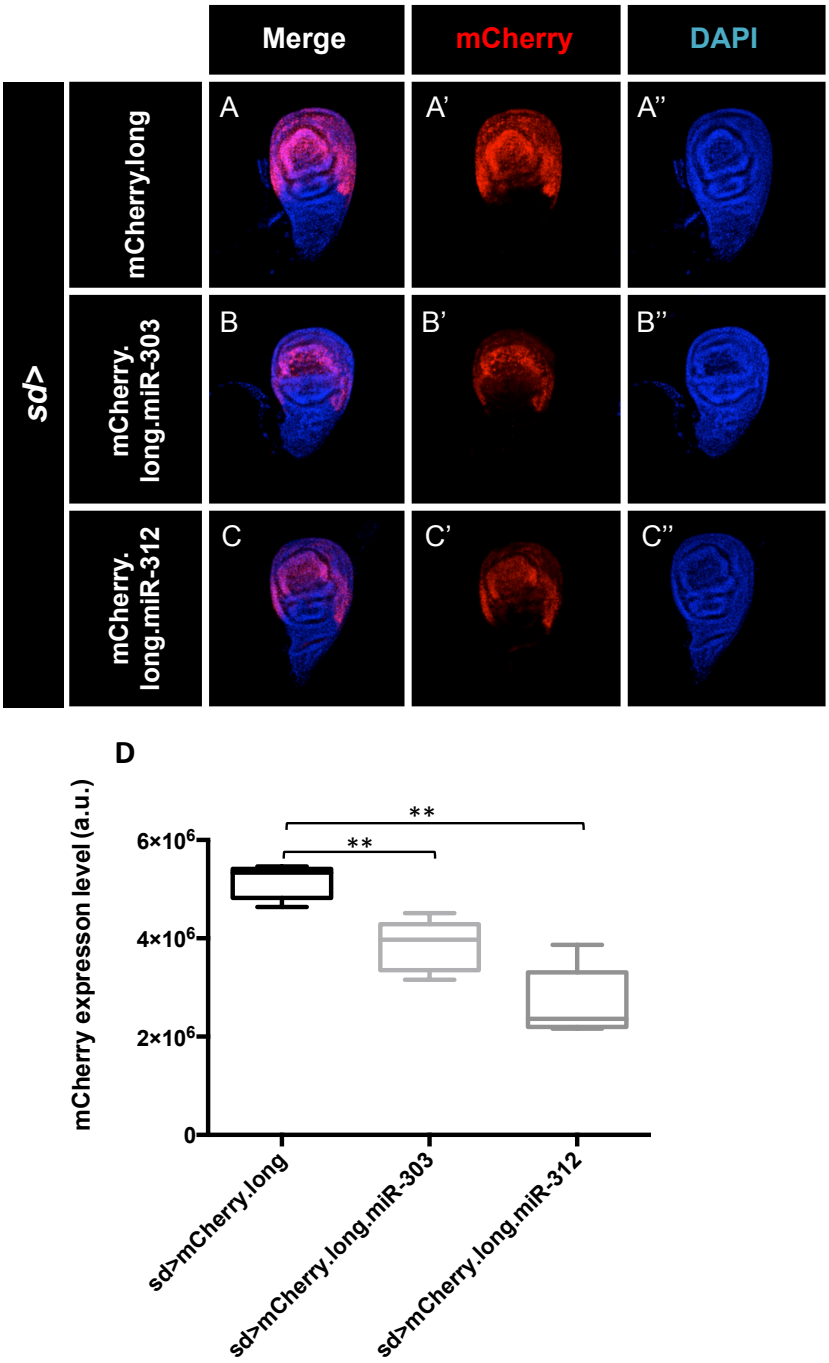


Figure 4.11 Validation of physical interactions between miRNAs and *Ubx* 3' UTR (Part I).

The *Ubx* long 3' UTR was used to test the interactions between *Ubx* 3' UTR and miRNAs. (A-C) Merged images for mCherry and DAPI staining. (A'-C') Antibody staining for mCherry. (A''-C'') DAPI staining. (A-C'') Third instar larval haltere imaginal discs that carry UAS-mCherry *Ubx* long 3' UTR, *sd*-Gal4 and UAS-miRNA (linked to a miRNA in B-C''). (D) The x-axis indicates the genotype of the sample. The *y*-axis denotes mCherry expression calculated by arbitrary fluorescence intensity units (a.u.). An average of 5 haltere imaginal discs of each genotype was used in all analysis. The intensities of the staining were quantified by ImageJ, and statistical analysis were performed using a Mann-Whitney unpaired *t*-test; P<0.01 (**). *miR-303* and *miR-312* directly down-regulate *Ubx*-3'UTR expression in the *Drosophila* haltere imaginal disc.

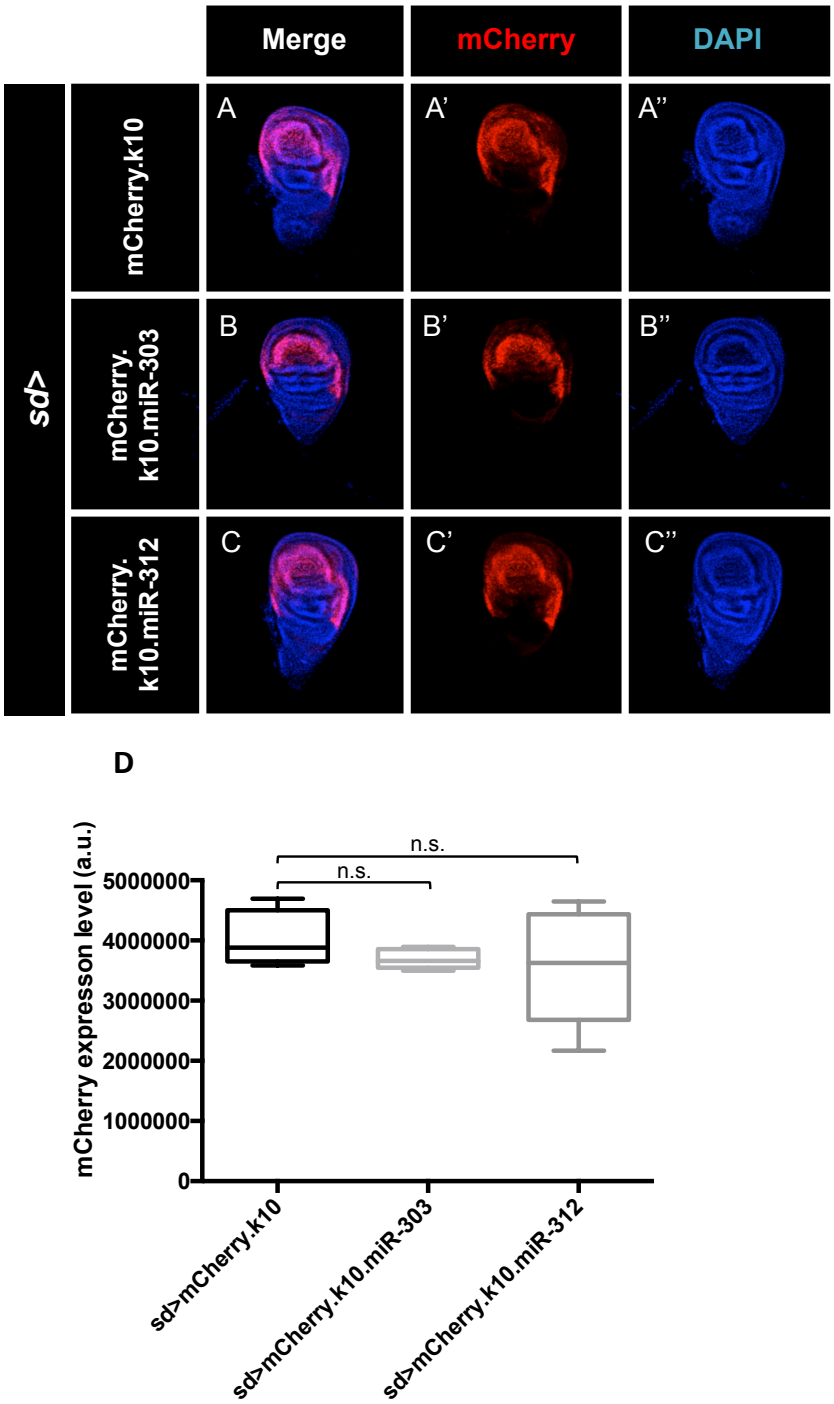


Figure 4.12 Validation of physical interactions between miRNAs and *K10* 3' UTR (Part II).

The *K10* 3' UTR was used as a control to make sure that the interactions observed between miRNAs and *Ubx* 3' UTR are not artificially affected. (A-C) Merged images for mCherry and DAPI staining. (A'-C') Antibody staining for mCherry. (A''-C'') DAPI staining. (A-C'') Third instar larval haltere imaginal discs that carry UAS-mCherry *K10* 3' UTR, *sd*-Gal4 and UAS-miRNA (linked to a miRNA in B-C''). (D) The x-axis indicates the genotype of the sample. The *y*-axis denotes mCherry expression. An average of 5 haltere imaginal discs of each genotype was used in all analysis. The intensities of the staining were quantified by ImageJ, and statistical analysis were performed using a Mann-Whitney unpaired *t*-test; P>0.05 (non-significant; n.s.). *miR-303* and *miR-312* do not target *K10* 3' UTR in the *Drosophila* haltere imaginal disc.

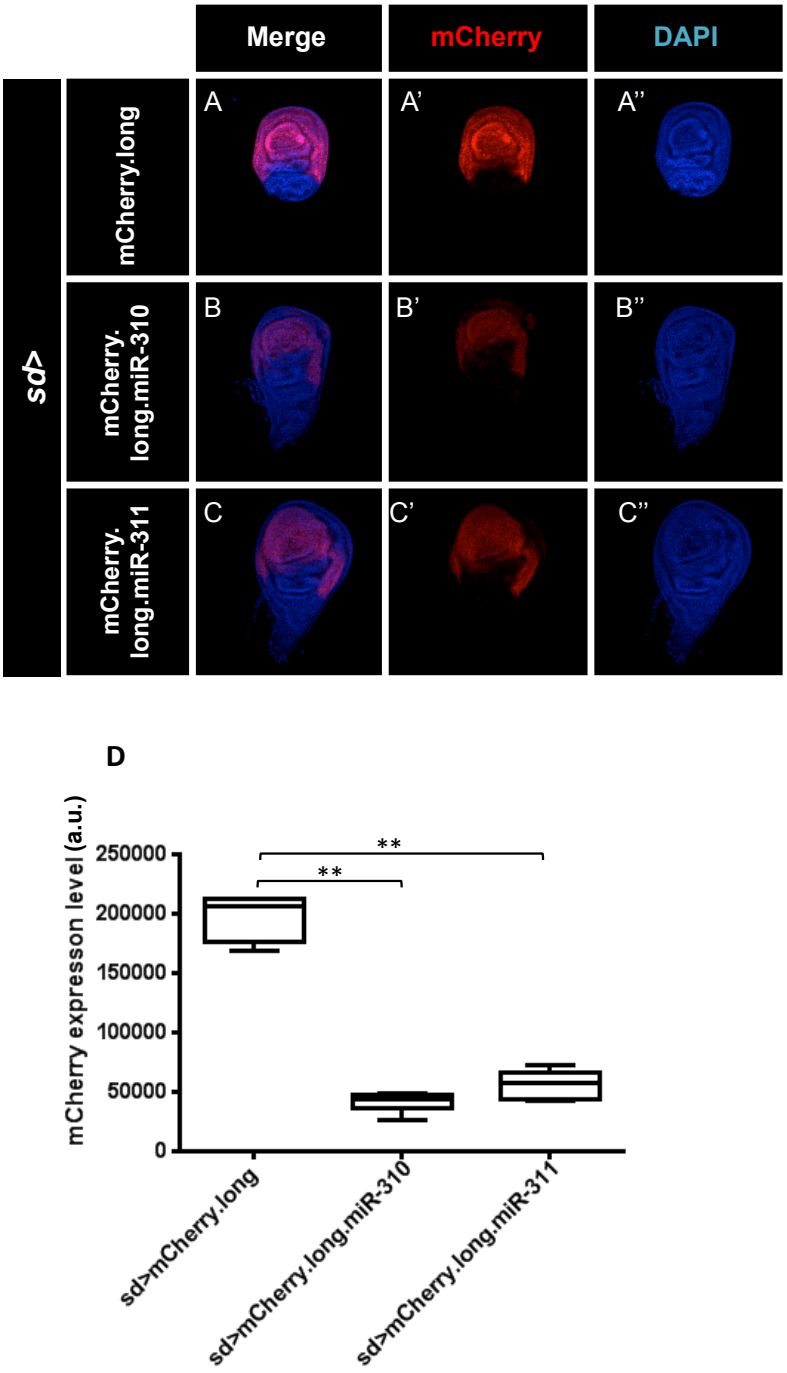


Figure 4.13 Validation of physical interactions between miRNAs and *Ubx* 3' UTRs (Part III).

The *Ubx* long 3' UTR was used to test the interactions between *Ubx* 3' UTR and miRNAs. (A-C) Merged images for mCherry and DAPI staining. (A'-C') Antibody staining for mCherry. (A''-C'') DAPI staining. (A-C'') Third instar larval haltere imaginal discs that carry UAS-mCherry *Ubx* long 3' UTR, *sd*-Gal4 and UAS-miRNA (linked to a miRNA in B-C''). (D) The x-axis indicates the genotype of the sample. The *y*-axis denotes mCherry expression. An average of 5 haltere imaginal discs of each genetype was used in all analysis. The intensities of the staining were quantified by ImageJ, and statistical analysis were performed using a Mann-Whitney unpaired *t*-test; P<0.01 (**). *miR-310* and *miR-311* directly down-regulate *Ubx*-3'UTR expression in the developing *Drosophila* haltere imaginal disc.

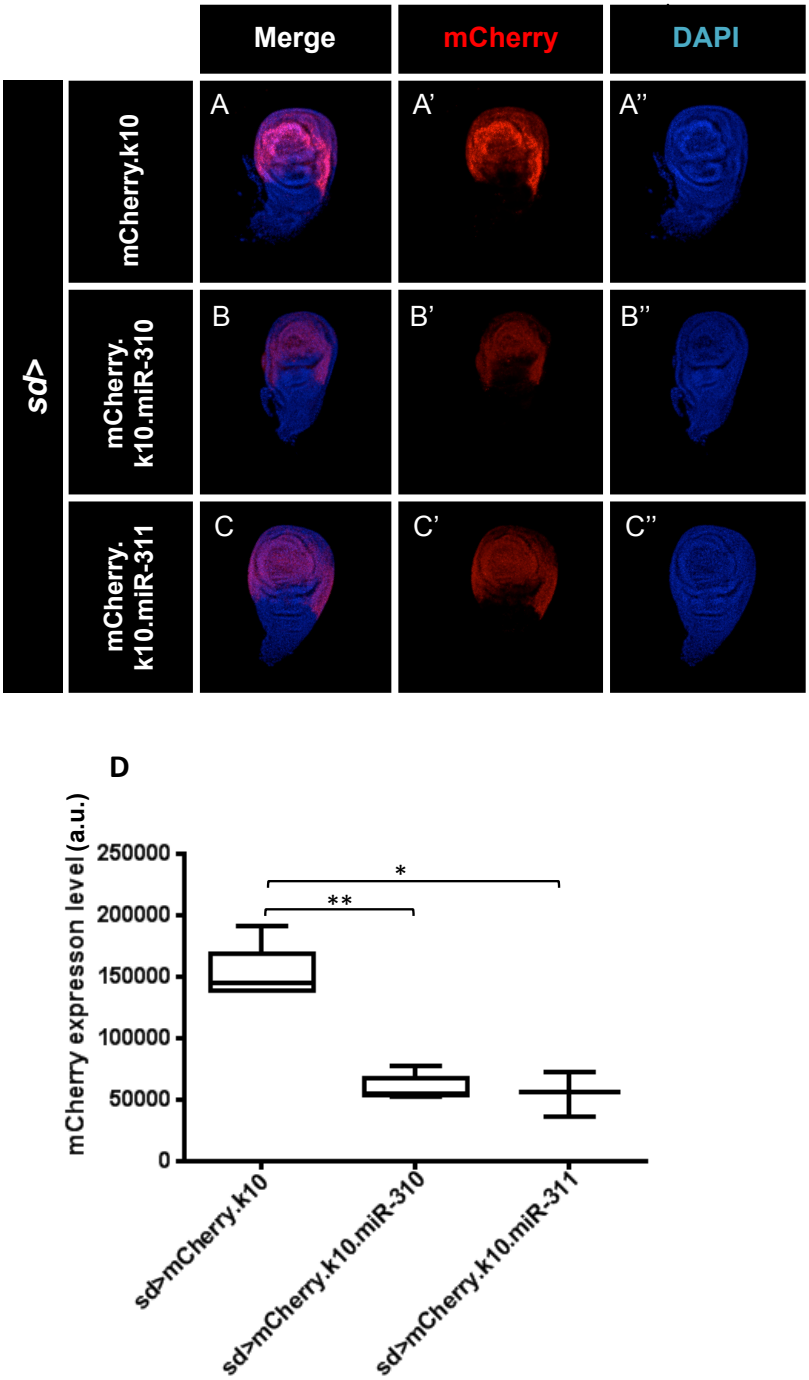


Figure 4.14 Validation of physical interactions between miRNAs and *K10* 3' UTR (Part IV).

The *K10* 3' UTR was used as a control to make sure that the interactions observed between miRNAs and *Ubx* 3' UTR are not artificially affected. (A-C) Merged images for mCherry and DAPI staining. (A'-C') Antibody staining for mCherry. (A''-C'') DAPI staining. (A-C'') Third instar larval haltere imaginal discs that carry UAS-mCherry *K10* 3' UTRs, *sd*-Gal4 and UAS-miRNA (linked to a miRNA in B-C''). (D) The x-axis indicates the genotype of the sample. The *y*-axis denotes mCherry expression. An average of 5 haltere imaginal discs of each genotype was used in all analysis. The intensities of the staining were quantified by ImageJ, and statistical analysis were performed using a Mann-Whitney unpaired *t*-test; $P<0.05$ (*), $P<0.01$ (**). *miR-310* and *miR-311* directly target *K10* 3' UTR in the *Drosophila* haltere imaginal disc.

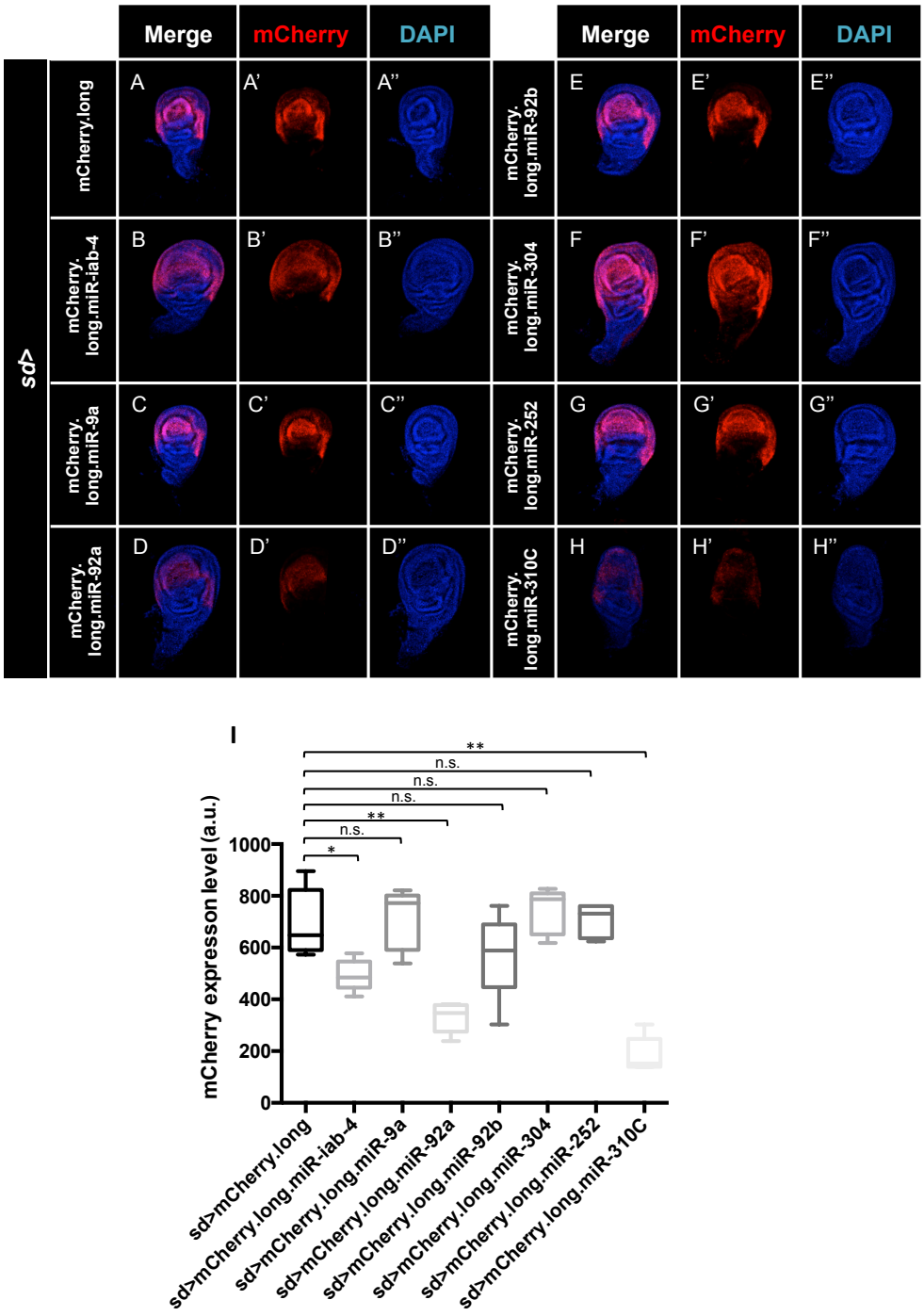


Figure 4.15 Validation of physical interactions between miRNAs and *Ubx* 3' UTR (Part V).

The *Ubx* long 3' UTR was used to test the interactions between *Ubx* 3' UTR and miRNAs. (A-H) Merged images for mCherry and DAPI staining. (A'-H') Antibody staining for mCherry. (A''-H'') DAPI staining. (A-H'') Third instar larval haltere imaginal discs that carry UAS-mCherry *Ubx* long 3' UTR, *sd*-Gal4 and UAS-miRNA (linked to a miRNA in B-H''). (I) The x-axis indicates the genotype of the sample. The *y*-axis denotes mCherry expression. An average of 5 haltere imaginal discs of each genotype was used in all analysis. The intensities of the staining were quantified by ImageJ, and statistical analysis were performed using a Mann-Whitney unpaired *t*-test; P>0.05 (non-significant; n.s.), P<0.05 (*), P<0.01 (**). *miR-iab-4*, *miR-92a* and *miR-310C* directly down-regulate *Ubx* 3'UTR expression in *Drosophila* haltere imaginal discs.

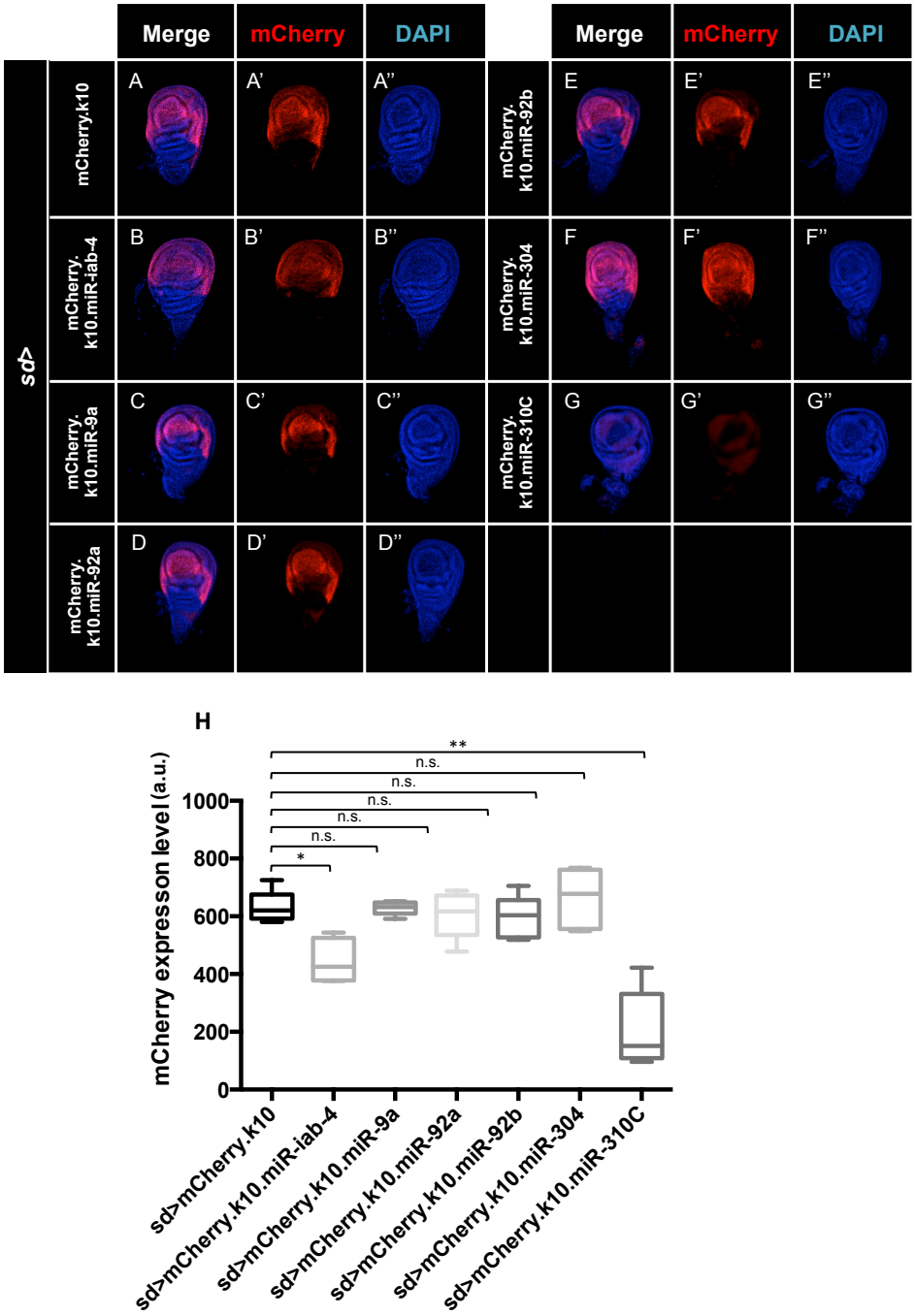


Figure 4.16 Validation of physical interactions between miRNAs and *K10* 3' UTR (Part VI).

The *K10* 3' UTR was used as a control to make sure that the interactions observed between miRNAs and *Ubx* 3' UTR are not artificially affected. (A-G) Merged images for mCherry and DAPI staining. (A'-G') Antibody staining for mCherry. (A''-G'') DAPI staining. (A-G'') Third instar larval haltere imaginal discs that carry UAS-mCherry *K10* 3' UTR, sd-Gal4 and UAS-miRNA (linked to a miRNA in B-G''). (H) The x-axis indicates the genotype of the sample. The *y*-axis denotes mCherry expression. An average of 5 haltere imaginal discs of each genotype was used in all analysis. The intensities of the staining were quantified by ImageJ, and statistical analysis were performed using a Mann-Whitney unpaired *t*-test; P>0.05 (non-significant; n.s.), P<0.05 (*), P<0.01 (**). *miR-iab-4* and *miR-310C* are directly targeting *K10* 3' UTR in the *Drosophila* haltere imaginal disc.

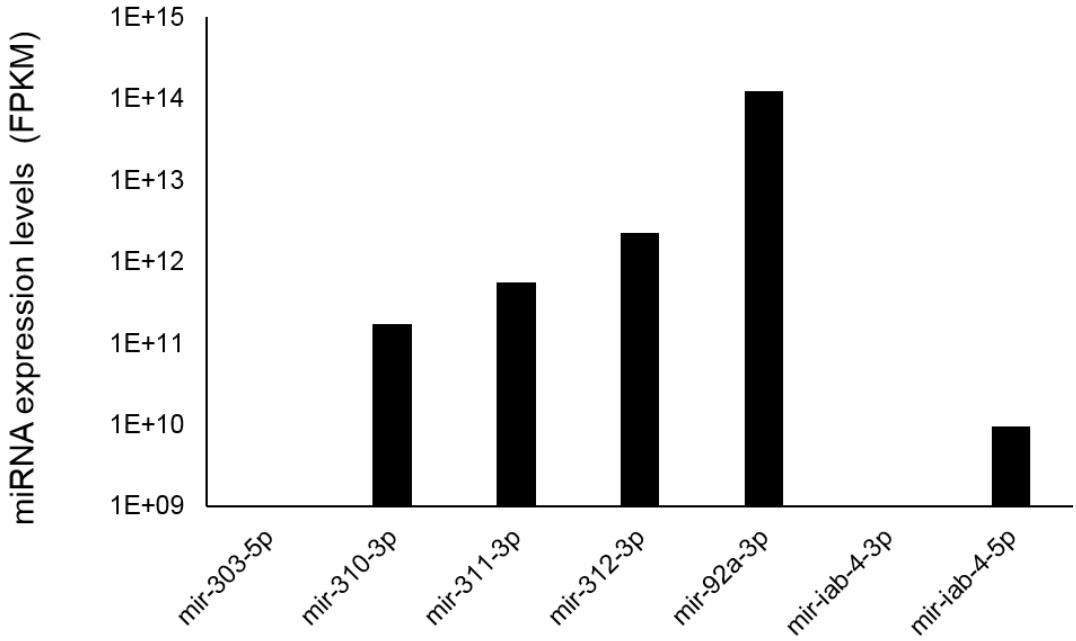


Figure 4.17 miRNA expression levels in white prepupal haltere imaginal discs by RNA-seq.

Graph showing the endogenous expression levels of miRNAs in white prepupal haltere imaginal discs by RNA-seq that appear to directly regulate *Ubx* 3'UTR expression when ectopically expressed **(collected by our previous lab member Dr. Richard Kaschula)**. The miRNA expression level was calculated by FPKM.

4.2.7 Experimental validation of the bioinformatic programs by the genetic screen results.

By investigating all ten miRNAs in Class I by experimental approaches, I found that 6 miRNAs directly target *Ubx* 3' UTR. When compared to this experimental process, bioinformatic approaches are a much more convenient way to find miRNAs that target a particular 3' UTR. However, different bioinformatic prediction tools have varying degrees of accuracy. Here, as a final note, I compared the experimentally-generated *Ubx* miRNA targeting data with bioinformatic predictions from a number of alternative softwares, in order to both understand which softwares perform best, as well as to find the most important parameters for miRNA targeting predictions.

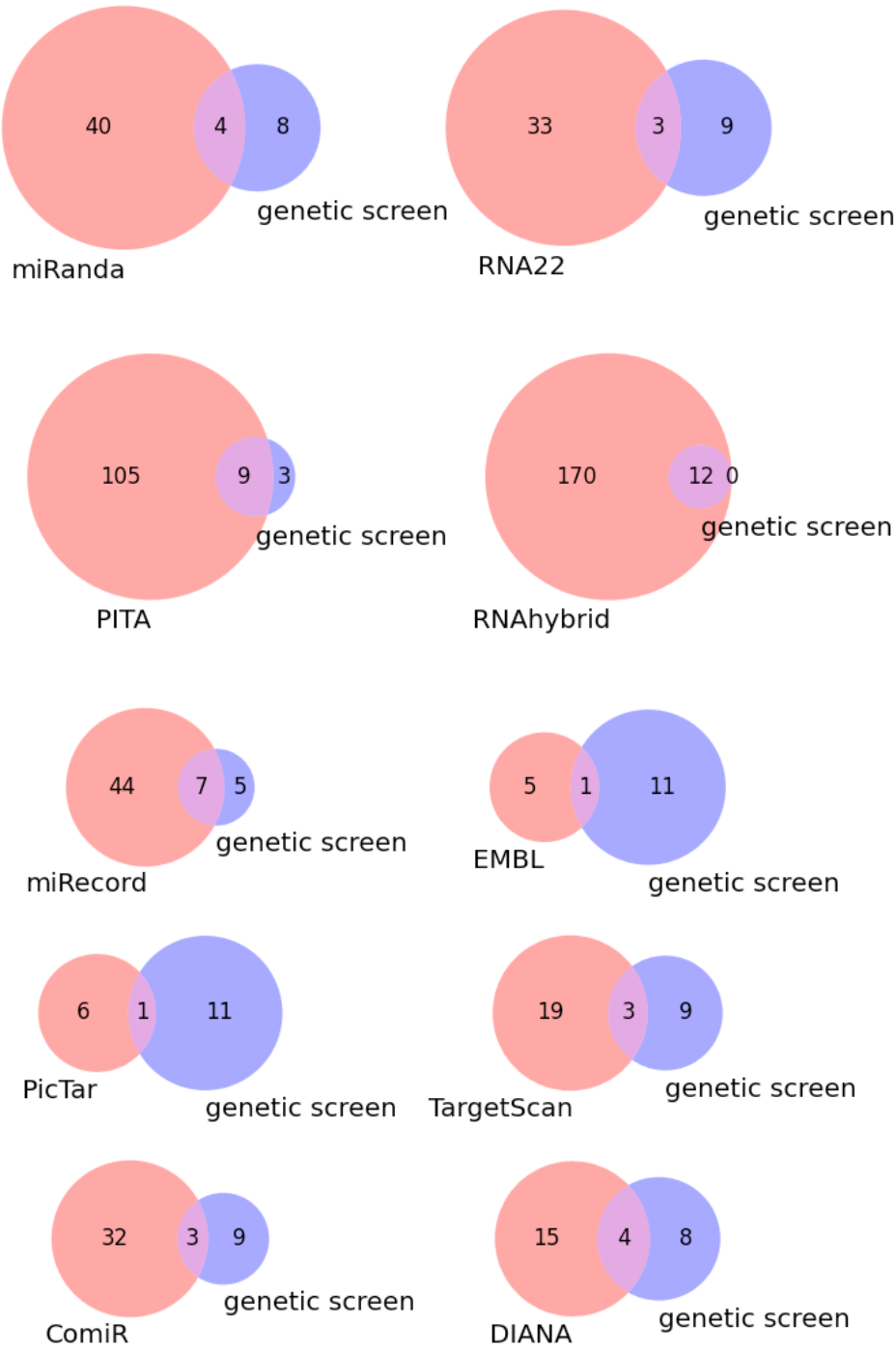


Figure 4.18 Overlap of miRNA target prediction results and genetic screen.

Red circle refers to mature miRNAs predicted to target *Ubx* by different softwares. Blue circles indicate miRNAs shown to produce *Ubx* mutant phenotypes by overexpression in haltere imaginal discs in the genetic screen (*miR-303-3p*, *miR-303-5p*, *miR-310-3p*, *miR-310-5p*, *miR-311-3p*, *miR-311-5p*, *miR-312-3p*, *miR-312-5p*, *miR-92a-3p, miR-92a-5p, miR-iab-4-3p* and *miR-iab-4-5p*). Pink area specifies miRNAs both predicted by programs and shown to elicit *Ubx*-like mutant phenotypes in our genetic screen.

The six miRNAs, *miR-303*, *miR-310*, *miR-311*, *miR-312*, *miR-92a* and *miR-iab-4*, were shown to directly target *Ubx* 3' UTRs in *vivo*. Here, I used these experimental data to validate the performance of miRNA target prediction programs. Figure 4.18 shows the overlap in *Ubx* targeting miRNAs between bioinformatic predictions and genetic screen.

Among the four programs that have the same number of inputted miRNAs and a similar annotated *Ubx* 3' UTR, RNAhybrid has 12 miRNAs overlap with our genetic screen, the highest commonality between miRNA target predictions and our experimental results. However, RNAhybrid also displays the highest false positive rate (Figure 4.19). Briefly, PITA predictions recover 9 miRNAs that overlap with our genetic screen results, while miRanda recovers 4 miRNAs, thus showing the second and third highest sensitivity, respectively, while their false positive rates being lower than RNAhybrid. RNA22 displays the lowest sensitivity of all tested softwares, but also the lowest false positive rate. This is expected as increased sensitivity always occurs at the cost of specificity.

I examined the common miRNAs between these four prediction programs and found 13 miRNAs that were predicted to target *Ubx* 3' UTRs (Figure 4.20). However, among them, only *miR-312-3p* directly targets *Ubx* 3' UTRs according to the experimental result. As such, the false positive rate for the combined prediction is 92.3% (12 out of 13). I also miss five miRNAs (*miR-303*, *miR-310*, *miR-311*, *miR-92a* and *miR-iab-4*) that directly target *Ubx* 3' UTRs. As previously stated, the main problem for miRNA targeting prediction programs is confirmed as the high false positive rate.

For the other six programs, the false positive rates are still very high. miRecord shows the best performance, predicting targets that include all genuine miRNAs that were experimentally shown to directly regulate *Ubx* expression. miRecord collates the predictions of 12 different miRNA target prediction softwares, and it is perhaps not surprising that it recovers all six *Ubx*-targeting miRNAs, showing the advantage in combining different miRNA prediction tools. However, the false positive rate for miRecord is high (Figure 4.19). DIANA, TargetScan and ComiR all have 4, 3 and 3 common hits between bioinformatics predictions and genetic

screens, respectively. Thus, I suggest that the prediction efficiency of these four programs is not ideal for *Drosophila* miRNA target prediction. EMBL and PicTar had the lowest common hits with only 1 miRNAs in both cases. However, the number of miRNAs considered in the prediction is limited, and I cannot, therefore, determine the performance of these programs. Also, since the algorithm for these two programs is not openly available (Table 1.2) and thus not available for peer scrutiny, these programs are not recommended for miRNA prediction. Based on this analysis, I can state that no miRNA target prediction program achieves full efficacy and conclude that combining results from different miRNA prediction softwares should be a way to improve prediction sensitivity although at the cost of specificity.

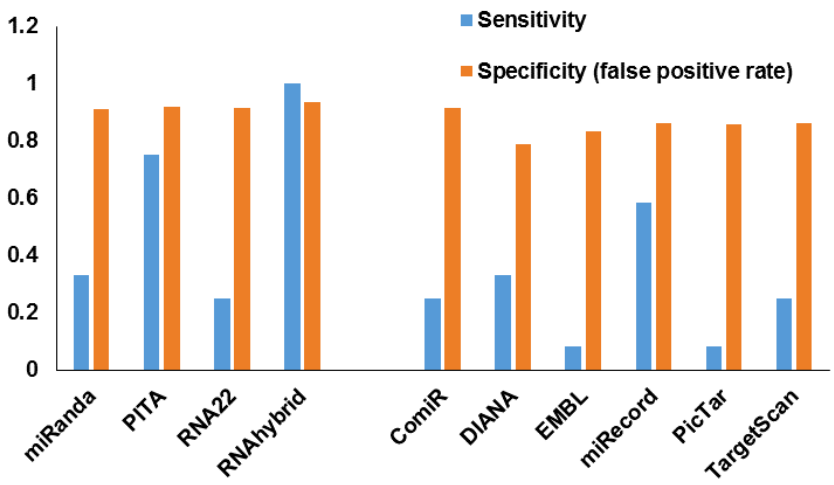


Figure 4.19 Specificity and sensitivity of miRNA target prediction programs in relation to the results of our genetic screen.

Specificity (specificity=predicted targeting of *Ubx* but not experimentally observed targeting for *Ubx* / all predicted targets for *Ubx*) and sensitivity (sensitivity= predicted and experimentally validated targeting of *Ubx* / all experimentally observed targeting of *Ubx*) of each miRNA target prediction program. Blue columns refer to the sensitivity of each program and orange columns indicate its specificity.

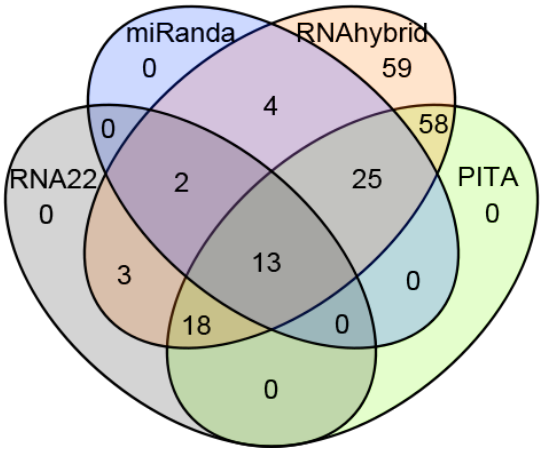


Figure 4.20 The common miRNAs predicted target *Ubx* 3' UTR by four bioinformatic programs.

Venn diagrams representing the common miRNAs between these four prediction programs, miRanda, PITA, RNAhybrid and RNA22, and found 13 miRNAs (*dme-mir-1003-5p*, *dme-mir-12-5p*, *dme-mir-137-5p*, *dme-mir-184-3p*, *dme-mir-279-5p*, *dme-mir-280-5p*, *dme-mir-309-3p*, *dme-mir-312-3p*, *dme-mir-318-3p*, *dme-mir-375-3p*, *dme-mir-932-3p*, *dme-mir-966-5p*, *dme-mir-987-5p*) are predicted to target *Ubx* 3' UTR.

4.2.8 Genuine miRNA target sites

Each miRNA transgenic fly contains a miRNA precursor sequence that has the ability to produce either one to two mature miRNAs. My screen results do not provide sufficient information to determine the functional effects of individual mature miRNAs. In order to understand which predicted miRNA target sites in the *Ubx* 3' UTRs are indeed genuine for the six miRNAs above, I need to know which mature miRNAs are functional. To have an overview of which arm of the primary miRNA transcript dominates the regulation of *Ubx* 3' UTRs, I looked into the miRNA prediction results of RNAhybrid, PITA, miRanda and RNA22 for these six miRNAs (Table 4.1).

RNAhybrid results are indistinguishable between both mature forms for all miRNAs. The result for PITA predictions is different, as only the following mature forms *miR-303-5p*, *miR-311-3p* and *miR-312-3p* are predicted to target *Ubx* 3' UTR, in addition to mature miRNAs stemming from both arms of the *miR-310*, *miR-92a* and *miR-iab-4* precursors. miRanda only predicts *Ubx* 3' UTR targeted by *miR-303-5p*, *miR-312-3p*, *miR-iab-4-5p* and *miR-iab-4-3p*. RNA22 provides the smallest amount of predicted targets, with only *miR-310-5p*,

miR-312-3p and *miR-92a-5p* as high-confidence regulators of *Ubx* 3' UTRs. There are lots of differences among miRNA targeting prediction results of different programs. However, which are the genuine target sites for the six miRNAs that directly regulate *Ubx* expression?

Table 4.1 miRNA target prediction for the miRNAs directly interact with *Ubx* 3' UTR.

	miR-303	*miR-310*	*miR-311*	*miR-312*	*miR-92a*	*miR-iab-4*
miRanda	5p	N	N	3p	N	5p3p
PITA	5p	5p3p	3p	3p	5p3p	5p3p
RNA22	N	5p	N	3p	5p	N
RNAhybrid	5p3p	5p3p	5p3p	5p3p	5p3p	5p3p

3p5p indicates both arms of this miRNA predicted target *Ubx*. 3p indicates 3' arm predicted target *Ubx*. 5p means 5' arm predicted target *Ubx*. N indicates this miRNA does not predicted target *Ubx*.

To address above questions, I investigated the predicted targets for all six miRNAs that experimentally show to target *Ubx* 3' UTRs (Figure 4.21). Interestingly, all six miRNAs have target sites on both short and long *Ubx* 3' UTRs. It has been previously reported that the alternative processing of *Ubx* 3' UTRs leads to the lack or addition of potential miRNA target sites that can modulate differential interactions between particular miRNA species and subsets of mRNAs isoforms (Bartel & Chen 2004). Thus, I suggest that target sites on both short and long *Ubx* 3' UTR may be used to modulate interactions between miRNAs and *Ubx* isoforms. To confirm this hypothesis, I need to validate which miRNA target sites are not spurious i.e. biologically active. I suggest different experimental approaches to test the biological activity of individual miRNA target sites. Firstly, I could employ an *in vitro* approach that cotransfects constructs of mature miRNA and *Ubx* gene with a specific modified 3' UTR that individual mature miRNAs predicted target site is mutated, into *Drosophila* S2 cells; I could then detect *Ubx* expression level changes (compared to S2 cell controls that are with constructs of mature miRNA and *Ubx* gene with normal *Ubx* 3' UTR), confirming the biological activity of an individual mature miRNA-3' UTR target site pair. Additionally, I can use an *in vivo* approach uncover genuine miRNA target sites in the *Ubx* 3' UTR. I can create a transgenic *Drosophila* stock where an individual target site in the *Ubx*

3’ UTR is mutated (e.g. using CRISPR), measuring the changes in Ubx expression, with the additional possibility of acquiring knowledge about the biological role of this individual target site through phenotypic analysis.

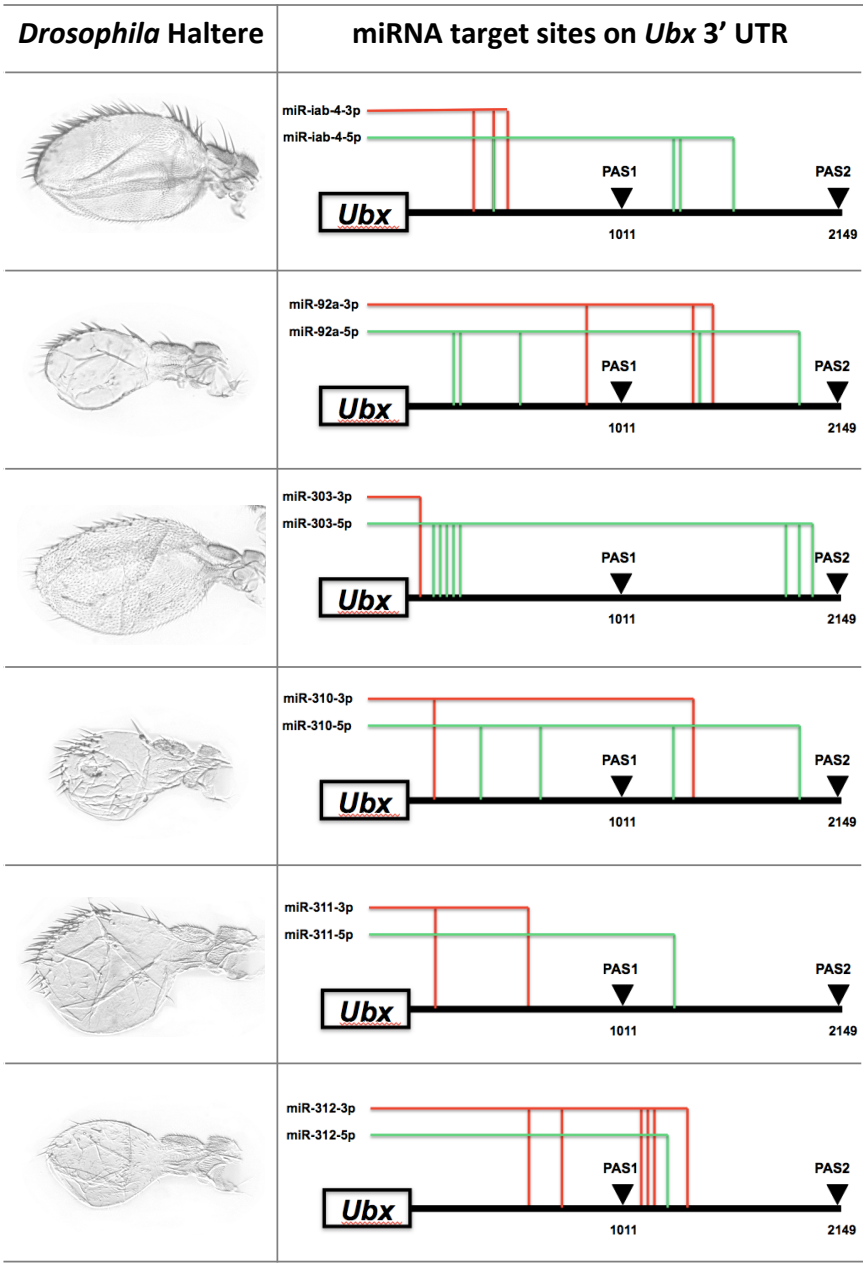


Figure 4.21 Diagrams depicting miRNA and mRNA interactions for potentially genuine target sites of the miRNAs that directly regulate *Ubx* expression. The black line denotes the full length *Ubx* long 3’ UTR. PAS 1 and PAS 2 indicate first polyadenylation site and second polyadenylation site, respectively. PAS 1 locates at 1011 nucleotides into the 3’ UTR sequence; the full length *Ubx* long 3’ UTR is 2149 nucleotides in length. Green lines refer to the target sites of 5’ arm of the mature miRNAs, while red lines depict target sites for the 3’ arm of mature miRNAs. All six miRNA have various target sites on both short and long *Ubx* 3’ UTRs, which may be used to reduce or add miRNA target sites in order to modulate interactions between miRNAs and *Ubx* isoforms.

4.3 Discussion

In this chapter, I expanded the work described in Chapter 3 by validating my genetic screen results combining experimental and bioinformatic strategies to generate hypotheses about genuine target sites for six miRNAs that directly interact with *Ubx* 3' UTR.

First, by combining the cuticle preparation and SEM, I observed haltere morphological changes after ectopic expression of miRNAs in high detail. The miRNAs from Class I show various haltere morphological changes. I suggest that the severity of the haltere morphological change is positively correlated with the levels of *Ubx* expression repression. Indeed, by using the immunohistochemistry, I examined Ubx protein expression levels in the haltere imaginal discs and found that nine miRNAs (90%) in Class I are able to regulate *Ubx* expression in the haltere imaginal discs when ectopically expressed, and the relative reduction in Ubx protein expression for each miRNA is consistent with the severity of the morphological change observed, with the exception of *miR-303*. I suggest that a possible cause for this deviation to the rule is, that in addition to regulating *Ubx* expression in haltere imaginal discs, *miR-303* may also regulate *Ubx*-targeted genes that control the haltere morphology (Weatherbee et al. 1998). Thus, even though *miR-303* repress *Ubx* expression in a less-efficient manner, the regulation of *Ubx* target genes by *miR-303* may cause an additive effect in terms of haltere morphological change, leading to similar functional results to those of miRNAs solely targeting *Ubx* but with higher efficiency.

More Interestingly, miRNAs from the *miR-92* family show different levels of homeotic transformation upon ectopic expression in the haltere imaginal disc, suggesting that besides seed sequence, other factors may affect the targeting efficiency of miRNAs in the *miR-92* family. It has been shown that the miRNA 3' end determines target specificity within a miRNA family (Brennecke et al. 2005). By using phylogenetic analysis and sequence alignments for mature miRNA sequences, I found that miRNAs that play functionally similar roles in post-

transcriptional regulation tend to cluster together (i.e. have similar sequences). The alignment of mature miRNA sequence shows that the t1A position may play a very important role in miRNA-mediated repression, consistent with previous reports that t1As are conserved in many mammalian miRNA target sites and in addition enhance miRNA-mediated repression (Bartel 2009; Lewis et al. 2005). It has been suggested that miRNA abundance affects the choice of binding sites in the miRNA-mRNA interaction (Coronnello et al. 2012). Thus, more abundant miRNAs in the same miRNA family may cause more severe effects after ectopic expression. However, I did not find any correlation between the endogenous expression levels of miRNAs in the haltere imaginal disc and the differential severity of morphological changes among miRNAs in *miR-92* family after ectopic haltere disc expression. I suggest that the differential number and target affinity of miRNA target sites for each miRNA in the *Ubx* 3' UTR could be additional factors that explain these results.

In addition to homeotic transformations, miRNAs in *miR-92* family also show loss of the hairs that cover the haltere surface. This feature is consistent with previous reports that ectopic expression of *miR-92a* and its seed relatives have the unique ability to trigger trichome loss (Bejarano et al. 2012; Schertel et al. 2012).

Second, I employed a genetic approach to assess postulated physical interactions between the *Ubx* 3' UTR and different miRNAs and found that six miRNAs in Class I directly interact with *Ubx* 3' UTRs. Five of these miRNAs are endogenously expressed in the haltere, indicating potential biological roles for these miRNAs in haltere development. Among them, only *miR-iab-4* is contained within the *Hox* cluster.

Additionally, I used the six miRNAs to examine the four bioinformatics prediction programs and found that no miRNA target prediction program achieves full efficacy and conclude that combining results from different miRNA prediction softwares should be a way to improve prediction sensitivity although at the cost of specificity.

At the end of this chapter, I combined the results from bioinformatic predictions and experimental genetic screen to find genuine miRNA target sites in *Ubx* 3' UTRs. However, further experiments are needed to confirm the biological role of these target sites. If this strategy is successful, I suggest a streamlined approach to find miRNA target sites on the 3' UTR of a particular *Drosophila melanogaster* gene, involving (i) screening using a particular phenotype of a gene loss-of-function; (ii) using immunohistochemistry and reporter constructs to validate direct targeting; (iii) using bioinformatics programs to predict specific target sites; (iv) mutate putative target sites and use S2 cells or transgenic flies to validate the result.

In summary, I conclude that six miRNAs out of the 96 miRNAs screened, *miR-303*, *miR-310*, *miR-311*, *miR-312*, *miR-92a* and *miR-iab-4*, directly regulate *Ubx* expression in the haltere imaginal disc, indicating that the genetic screen is a valuable and efficient tool to test regulatory interactions between miRNAs and mRNAs molecules. Additionally, the successful conduction of this experiment provides a list of miRNAs regulating *Ubx* expression during haltere development. I hope to further investigate the biological function of these miRNAs in the future, and with this provide new insights into the study of miRNA targeting mechanisms.

CHAPTER 5 The regulatory roles of miRNAs within *Drosophila melanogaster* embryonic CNS.

5.1 Introduction

During development, differential gene expression leads to the specification of cellular identities (Alonso 2012; Orphanides & Reinberg 2002; Alberts et al. 2002). In this context, differentially expressed miRNAs contribute to cell specification by providing a unique biochemical environment for gene expression. This in turn refines the molecular responses that control cell differentiation and cell fate decisions on multiple tissues (Ivey & Srivastava 2010; Shenoy & Blelloch 2014). Thus, understanding the molecular mechanisms of miRNA regulation at a cellular level is crucial to study differential gene expression and function. In this chapter, I seek to extend my work into the cellular dimension, focusing on the roles played by miRNAs in the regulation of *Hox* gene expression within the CNS.

In the previous chapter, I found that nine miRNAs from the phenotypic Class I can regulate *Ubx* expression in the developing *Drosophila* haltere. However, do these miRNAs endogenously target *Ubx* during development? If so, can this regulation explain cellular-level variations in *Ubx* expression? Although previous research has shown that different levels of the gene regulation hierarchy control haltere morphology (Weatherbee et al. 1998), the molecular mechanisms involved in haltere formation and the cellular level regulation of the haltere are not fully understood. It is thus difficult to relate specific phenotypes to effects in specific cells. As such, the haltere is not a good system to study miRNA-*Ubx* target recognition at the cellular level.

For the purpose of studying the effects of miRNA-*Ubx* interactions at the cellular level, I focused on the CNS as it allows linking gene function to individually identifiable cells. By using the DiI cell-labelling technique, parental neuroblasts for almost all cell types in the *Drosophila* CNS have been identified (Bossing et al. 1996; Schmidt et al. 1997; Schmid et al. 1999). It is also possible to use

antibodies and enhancer trap lines to analyse gene expression related to the specification and differentiation of individual cells in the *Drosophila* CNS (Doe 1992; Beckervordersandforth et al. 2008; Ito et al. 1995). Recently, several groups generated and described the expression patterns of GAL4 lines active during different CNS developmental stages, including embryonic CNS expression of 5,000 GAL4 lines (Manning et al. 2012), larval CNS expression pattern of 6,650 GAL4 lines (Li et al. 2014), and expression patterns in the adult nervous system of 7,000 transgenic lines (Jenett et al. 2012), which greatly contribute to progress of the cell lineage analysis and gene manipulation in the *Drosophila* CNS. Thus, the *Drosophila* CNS is, so far, the best system to study the effects of miRNAs at the cellular level within the physiological context of development.

Ubx is expressed in the *Drosophila* CNS throughout development (Akam & Martinez-Arias 1985; Jarvis et al. 2012), which provides a good system for the cellular level interaction between miRNA and *Ubx*. Prokop and Technau reported that the activity of *Ubx* or *abd-A* determines the differentiation of the neuroblast NB1-1 lineage between thoracic and abdominal tagmata (Prokop & Technau 1994). *Ubx* is also necessary to activate *reaper*-dependent cell death of specific motorneurons in the *Drosophila* embryonic CNS (Rogulja-Ortmann et al. 2008). By manipulating *Ubx* expression in postembryonic CNS, Marin *et al.* found that *Ubx* determines the segment-specific features of neuron morphology and survival pattern (Marin et al. 2012).

Furthermore, several studies found that *miR-iab-4* and *miR-iab-8* regulate *Ubx* expression during *Drosophila* embryonic CNS development (Bender 2008; Thomsen et al. 2010). Meza-Sosa *et al.* reported that miRNAs play an important role during CNS development and are also involved in neuropathologies such as neurodegenerative diseases, developmental CNS disorders, and psychiatric disorders (Meza-Sosa et al. 2012). Although the *Drosophila* brain is far simpler than the human brain, the neurons as individual cells in the human brain and those in *Drosophila* are similar (Cohen 2010). Thus, studying the cellular level regulation of miRNA function in *Drosophila* can shed light to understand the mechanism of gene regulation in other systems, such as humans, which could

be ultimately useful for the development of disease diagnoses and therapies.

In this chapter, I explore the nature and effects of miRNA-based regulation of *Ubx* mRNAs at the cellular level within the *Drosophila* CNS. First, I analyse the miRNA expression levels for those miRNAs repressing *Ubx* expression in phenotypic Class I by using previously published miRNA sequencing data (Ruby et al. 2007). I found that, except for *miR-303* and *miR-304* that are barely expressed during embryogenesis, seven miRNAs out of nine are all expressed during embryo development. Secondly, by using fluorescent *in situ* hybridisation, I found that *miR-252* and *miR-92a* are expressed in a complementary manner with *Ubx* in the *Drosophila* CNS. Finally, I compared *Ubx* expression levels in the *Drosophila* CNS between wild-type and *miR-252* mutants in *repo*-marked glial cell lineages, and between wild-type and engrailed-marked *miR-92a* mutant embyos. I found that the expression level of *Ubx* increases in the glial cells of *miR-252* mutant CNS. However, I did not find any *Ubx* expression level change at all in *miR-92a* mutant CNS.

5.2 Results

5.2.1 Temporal and spatial expression during embryogenesis for miRNAs that have the ability to regulate *Ubx* expression in haltere.

The interaction between miRNA and their targets can only occur when these two molecular partners are co-expressed in time and space. To have an overview of potential interactions between miRNA and *Ubx* during *Drosophila* embryogenesis, I first analysed the temporal expression of those miRNAs that are proved to regulate *Ubx* expression in developing *Drosophila* haltere (Figure 4.7-4.9). By analysing the RNA-seq data from Ruby *et al.* (Ruby et al. 2007), I found that all nine miRNAs are expressed during embryogenesis (Table 5.1). However, the expression levels of *miR-303* and *miR-304* were extremely low, so I did not consider their roles during the *Drosophila* CNS development in the following analysis. In total, I found that 7 miRNAs are temporally co-expressed with *Ubx* during *Drosophila* embryogenesis.

Table 5.1 miRNA-*Ubx* temporal co-expression during *Drosophila* embryogenesis for those miRNAs that regulate *Ubx* expression in haltere by RNA-seq.

Data type	MiRNA	0-1 hr. embryos	2-6 hr. embryos	6-10 hr. embryos	12-24 hr embryos
Normalised	*miR-iab-4-5p*	2.28	13.74	53.65	20.71
	miR-iab-4-3p	16.12	37.27	36.87	4.65
	miR-252	0.83	0.14	3.24	11.71
	miR-303	0.00	0.00	0.00	1.34
	miR-304	0.00	1.83	1.23	9.78
	miR-92a	48.32	14.60	12.02	2.66
	miR-92b	38.93	17.77	16.34	1.87
	miR-310	62.29	14.63	5.45	1.06
	miR-311	49.89	22.28	6.90	1.60
	miR-312	44.52	25.27	15.72	2.28
Reads	*miR-iab-4-5p*	2.0	33.0	96.0	42.0
	miR-iab-4-3p	3.0	19.0	14.0	2.0
	miR-252	21.0	10.0	168.0	688.0
	miR-303	0.0	0.0	0.0	1.0
	miR-304	0.0	2.0	1.0	9.0
	miR-92a	660.0	546.0	335.0	84.0
	miR-92b	216.0	270.0	185.0	24.0
	miR-310	583.0	375.0	104.0	23.0
	miR-311	202.0	247.0	57.0	15.0
	miR-312	211.0	328.0	152.0	25.0

This set of miRNA sequencing data is from Ruby *et al.*, which sequenced 128 miRNAs throughout *Drosophila melanogaster* development, including the following embryogenesis stages: 0-1hr, 2-6hr, 6-10hr and 12-24hr after egg laying (Ruby et al. 2007). 'Reads' means the miRNAs reads in each sample while 'normalized' indicates the percentage of the reads in each sample dividing the total reads among all samples for each miRNA species. 'Reads' data gives us an absolute number of miRNAs in each sample, while 'normalized' shows us the relative miRNA abundance in each sample.

To validate the RNA-seq data and confirm the temporal expression pattern of the miRNAs of interest, I next selectively tested the temporal expression patterns for some of the miRNAs in Table 5.1 by semi-quantitative RT-PCR

(Figure 5.1A). The embryos were collected within a 24hr time-window after egg laying at two hour-intervals. I found that the expression patterns of both *miR-252* and *miR-92a* (Figure 5.1B) are coherent with the RNA-seq data (Table 5.1). This experiment not only confirms the results of the RNA-seq, but also provides a more detailed analysis of the temporal expression of these two miRNAs.

The APA of *Ubx* mainly produces two different 3' UTR isoforms: *Ubx* long 3' UTR (2149bp) and *Ubx* short 3' UTR (1011bp). An important study in our lab found that *Hox* mRNAs control their visibility to miRNA regulation by modulating the alternative 3' UTR processing system, and that the *Ubx* mRNA with the *Ubx* long 3' UTR is the dominant mRNA isoform during late neurogenesis (Thomsen et al. 2010). Therefore, if the miRNA expression pattern is similar to that of *Ubx* long 3' UTR, it is very likely there is a regulatory relationship between them in the developing *Drosophila* CNS.

miR-252 is highly expressed in late embryogenesis (8h-24h after egg laying), while *miR-92a* is highly expressed in early embryogenesis (0h-12h after egg laying) (Figure 5.1B). *miR-252* has a temporal expression pattern that is similar to the *Ubx* long 3' UTRs and thus antagonistic with the *Ubx* short 3' UTRs (Figure 5.1B) during neurogenesis (12h-24h after egg laying), which maybe indicative of a potential regulation of *Ubx* by *miR-252* during neurogenesis. In contrast, *miR-92a* expression decreases dramatically during neurogenesis. This pattern complements temporally with the expression pattern of *Ubx* long 3' UTRs (Figure 5.1B), reducing the possibility of a regulatory interaction between them. However, if the interaction between *miR-92a* and *Ubx* does exist, it may be limited to a small number of cells or occur with *Ubx* short 3' UTRs. Our results suggest that *miR-252* and *miR-92a* could regulate *Ubx* expression during embryonic neurogenesis.

The temporal expression of miRNAs only provides the information of the specific developmental timing in which a miRNA may regulate *Ubx* expression. However, it could be that miRNAs and *Ubx* mRNAs are highly expressed in the same developmental stage but in completely different cells, with no molecular interactions actually happening among them. Therefore, identification of the

spatial patterns of the two aforementioned miRNAs during embryogenesis is needed for probing the possibility of miRNA-*Ubx* interactions (Figure 5.2).

By using fluorescent *in situ* hybridisation, I found that *miR-252* (Figure 5.2A) and *miR-92a* (Figure 5.2B) are expressed in the *Drosophila* CNS. Interestingly, they are expressed in a "complementary" manner with Ubx (Figure 5.2A"', B"'), a pattern that is analysed in higher detail in Figure 5.4 (*miR-252*) and Figure 5.9 (*miR-92a*). However, *miR-92b* is expressed in the muscle (Figure 5.2C) rather than in the *Drosophila* CNS, which is consistent with the study by Chen *et al.* which reports that *miR-92b* is specifically expressed in heart and muscle, similarly to the expression pattern of Mef2 (Chen et al. 2012).

By using a *miR-310C* locus GAL4 insertion line, I detected that *miR-310C* are mainly expressed in the epidermis during early embryogenesis and in the brain during late embryogenesis (Figure 5.3A), being thus unlikely molecular partners of *Ubx* in the embryonic CNS (Figure 5.3A"'). I also looked for the expression patterns of *miR-310C* in the third instar larvae, finding that it is expressed in both the brain and posterior portion of the larval CNS (Figure 5.3B). Interestingly, *miR-310C* and Ubx express in a complementary manner in the posterior larval CNS (Figure 5.3B"'). This results show that *miR-310C* is possibly a regulator of *Ubx* expression in the developing larval CNS.

From the study of spatial expression patterns of miRNAs of interest (as determined by our genetic screen, see previous chapters), I found that *miR-252* and *miR-92a* could be potential regulators of *Ubx* expression in the *Drosophila* CNS during embryogenesis, while *miR-310C* may regulate *Ubx* expression during larval CNS development. Due to the time limitation, in the next sections, I will focus on the roles of *miR-252* and *miR-92a*, which could be in the possible regulation of *Ubx* expression during embryogenesis.

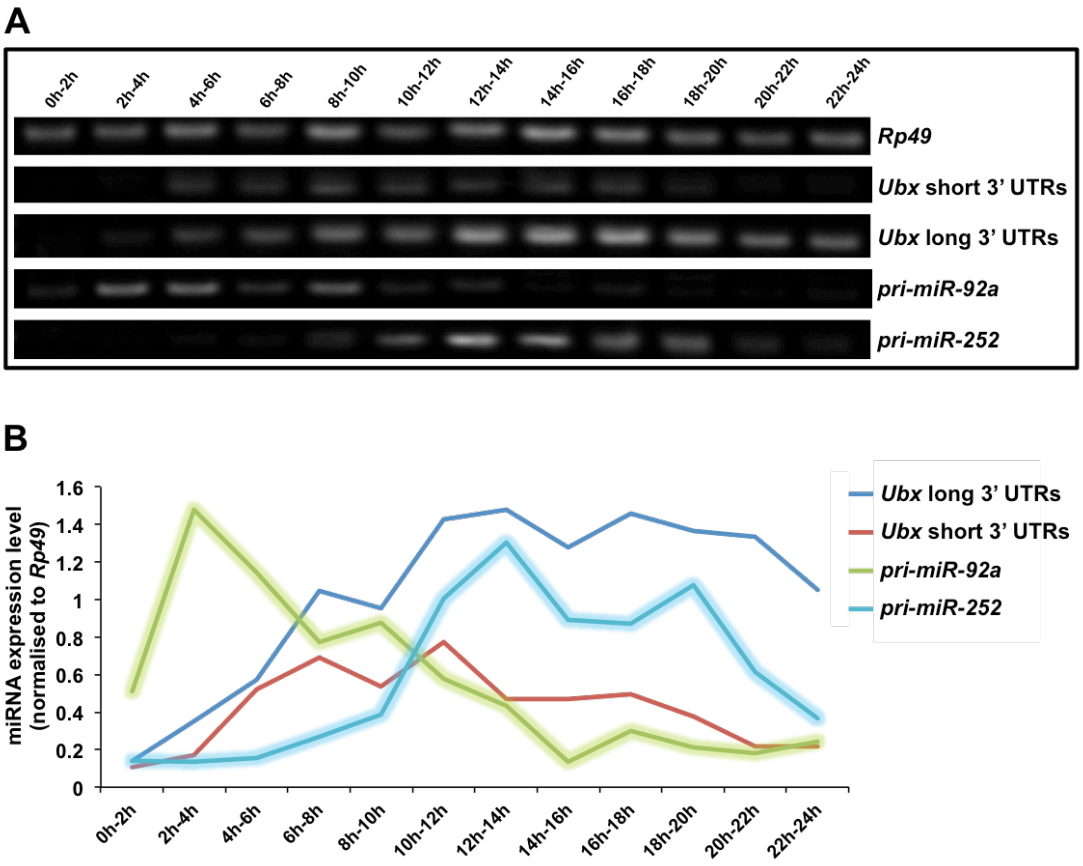


Figure 5.1 miRNA expression levels during embryogenesis by semi-quantitative RT-PCR.

(A) RT-PCR expression analysis of *miR-92a*, *miR-252*, *Ubx* long 3' UTRs and *Ubx* short 3' UTRs during embryogenesis with two hours sliding-window sample collection times from 0-24 hours after egg laying. The expression levels of Rp49 were used as a reference for all quantifications. (B) ImageJ quantification of the mRNA expression levels of each genes from the agarose electrophoresis gel in Figure 5.1A. The *miR-252* expression level is increased during late embryogenesis, while *miR-92a* is highly expressed in early embryogenesis, temporally decreasing as embryogenesis progresses. *miR-252* has a temporal expression pattern that is similar to the *Ubx* long 3' UTRs, while the pattern of *miR-92a* temporally contrasts with the expression pattern of *Ubx* long 3' UTRs.

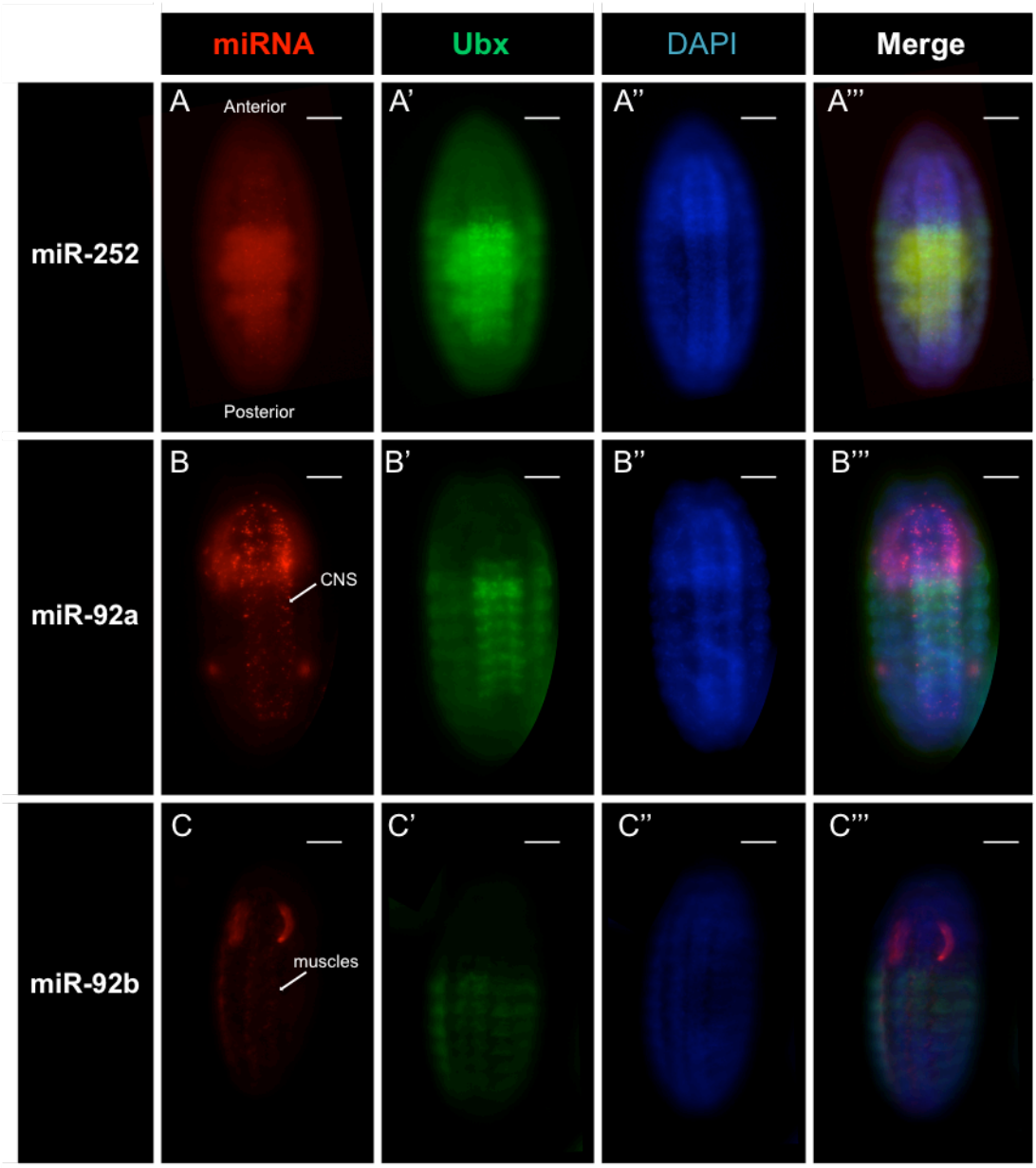


Figure 5.2 The spatial expression during embryogenesis for those miRNAs that regulate *Ubx* expression in haltere (Part I).

Fluorescent *in situ hybridis*ation for (A) *miR-252* in stage 16 embryo, (B) *miR-92a* in stage 16 embryo, and (C) *miR-92b* in stage 14 embryo. (A', B', C') Immunohistochemistry for Ubx. (A'', B'', C'') DAPI staining. (A''', B''', C''') Merged image. Both *miR-252* and *miR-92a* are expressed in the *Drosophila* CNS. The anterior is to the top. The arrows point to the CNS (B) and muscles (C) as indicated. Scale bar represents 40μm.

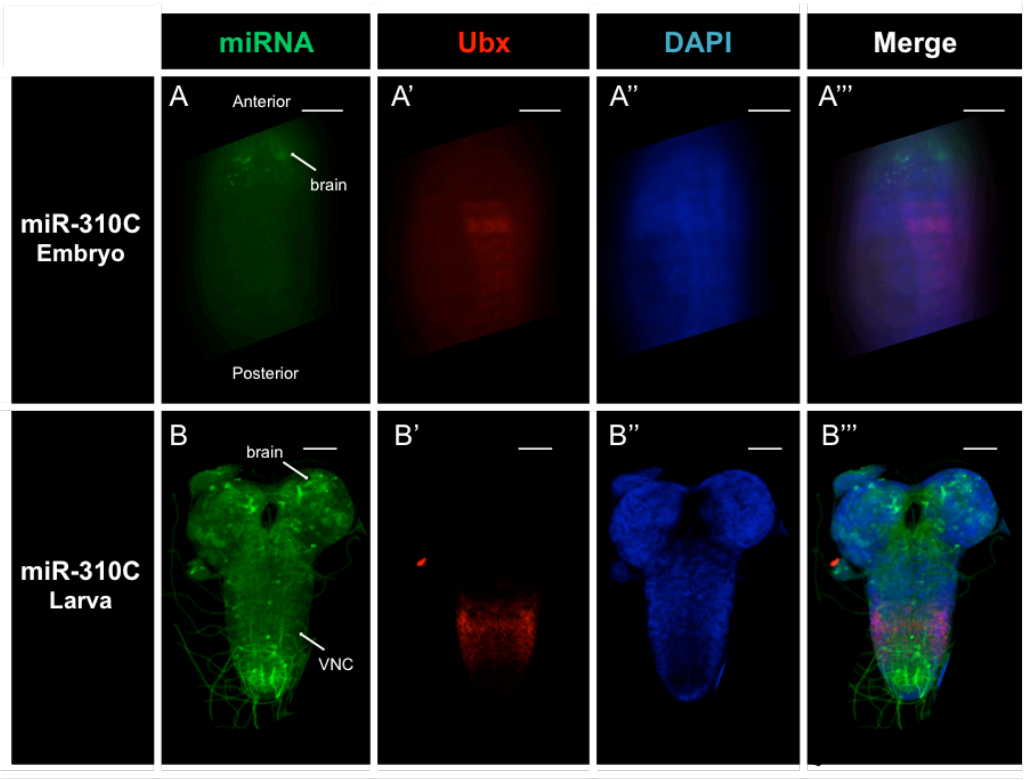


Figure 5.3 The spatial expression during embryogenesis for those miRNAs that regulate *Ubx* expression in haltere (Part II).

Immunohistochemistry for GFP expression driven by Gal4 insertion *P{GawB}NP5941* in the miR-310-313 (*miR-310C*) locus in *Drosophila* (A) stage 16 embryo and (B) third instar larva. CNS as shown (B) third instar larva CNS also places it in (B'). Immunohistochemistry for Ubx in *Drosophila* (A') embryo and (B') third instar larva. (A'', B'') DAPI staining. (A''', B''') Merged image. *miR-310C* is expressed in the brain in later stage embryos, and in both larval brain and posterior of larval CNS. The anterior is to the top. The arrows point to the brain (A, B) and VNC (B) as indicated. Scale bar represents 40μm.

5.2.2 The role of *miR-252* in the *Drosophila* CNS.

As *miR-252* is sufficient to repress *Ubx* expression in *Drosophila* haltere imaginal discs (Figure 4.8C', 4.9), and is also expressed in the *Drosophila* CNS where Ubx expression is present (Figure 5.2A), I next examined further the *miR-252* expression patterns in the *Drosophila* CNS to estimate the potential for molecular interactions between *miR-252* and *Ubx*. To this end, I combined a fluorescent *in situ* hybridisation approach for *miR-252* expression analysis with antibody stainings for the determination of Ubx expression. I found that *miR-252* and Ubx are expressed mostly in a complementary pattern in the *Drosophila* CNS (Figure 5.4B). As miRNA works to either repress gene translation or

trigger mRNA degradation (Filipowicz et al. 2008), it is very likely that the repressory effect of *miR-252* on *Ubx* leads to absence of Ubx expression level in specific CNS cell lineages. Therefore, I next asked whether *miR-252* does indeed regulate *Ubx* expression within the *Drosophila* CNS. And if this is so, how does this kind of regulatory interaction affect the *Drosophila* CNS development?

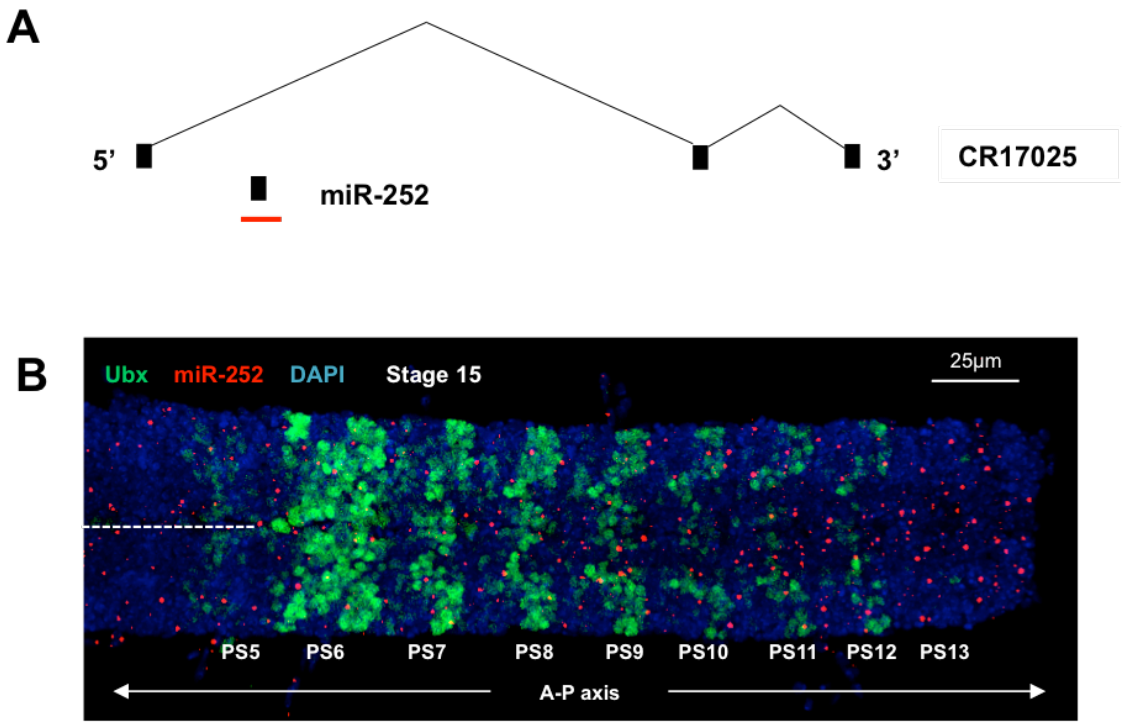


Figure 5.4 Spatial expression of *miR-252* within the *Drosophila* CNS.

(A) The diagram of the locus of *CR17025* denoting the probe designed for the detection of *miR-252* expression. The filled black squares are the exons for *miR-252* host gene *CR17025*, while the single lines that connect the squares are introns. The black square beneath *CR17025* gene is the position of the precursor *miR-252* in the intron region of *CR17025*. *miR-252* and its host gene *CR17025* are transcribed from left to right. The red line under the sequence of *miR-252* that spans the whole sequence of *miR-252* denoted the probe fragment used for *miR-252*. (B) Spatial expression of *miR-252* in stage 15 *Drosophila* CNS. *miR-252* is expressed in the whole CNS (Figure 5.2A). The red dots display *miR-252* expression. Green dots show Ubx expression in the CNS. DAPI-stained nuclei are coloured in blue. The *Drosophila* CNS anterior is to the left. *miR-252* and Ubx protein are expressed in a complementary manner within the developing *Drosophila* CNS. Scale bar represents 25μm.

5.2.2.1 *miR-252* is sufficient to repress Ubx expression in the *Drosophila* CNS.

miRNAs mainly act as repressors of gene expression. Therefore, if *miR-252* regulates *Ubx* expression in the *Drosophila* CNS, it should be sufficient to repress *Ubx* expression in this tissue. To test the inhibitory function of *miR-252*, I took advantage of the UAS-GAL4 system. Ectopic expression of *miR-252* was carried out using the neuronal-specific GAL4 driver *elav*-Gal4 and the effect on Ubx expression levels were evaluated (Figure 5.5B). In addition, the *elav-Gal4* line crossed with a *white* gene mutant fly was used as negative control, in order to exclude the possibility of *elav*-Gal4 and *white* gene mutants themselves having an effect on change of Ubx expression level. Additionally, I ectopically expressed *miR-iab-4* using the *elav*-Gal4 driver, and used the results as a positive control to test the efficiency of my genetic system, as it has been previously shown that *miR-iab-4* represses *Ubx* expression in the posterior *Drosophila* CNS (Bender 2008; Thomsen et al. 2010). Indeed, I found that Ubx expression level in CNS was reduced when I overexpressed *miR-252* in the *Drosophila* CNS, as compared to the negative control (Figure 5.5A). This result confirms that *miR-252* is sufficient to repress *Ubx* expression in the *Drosophila* CNS.

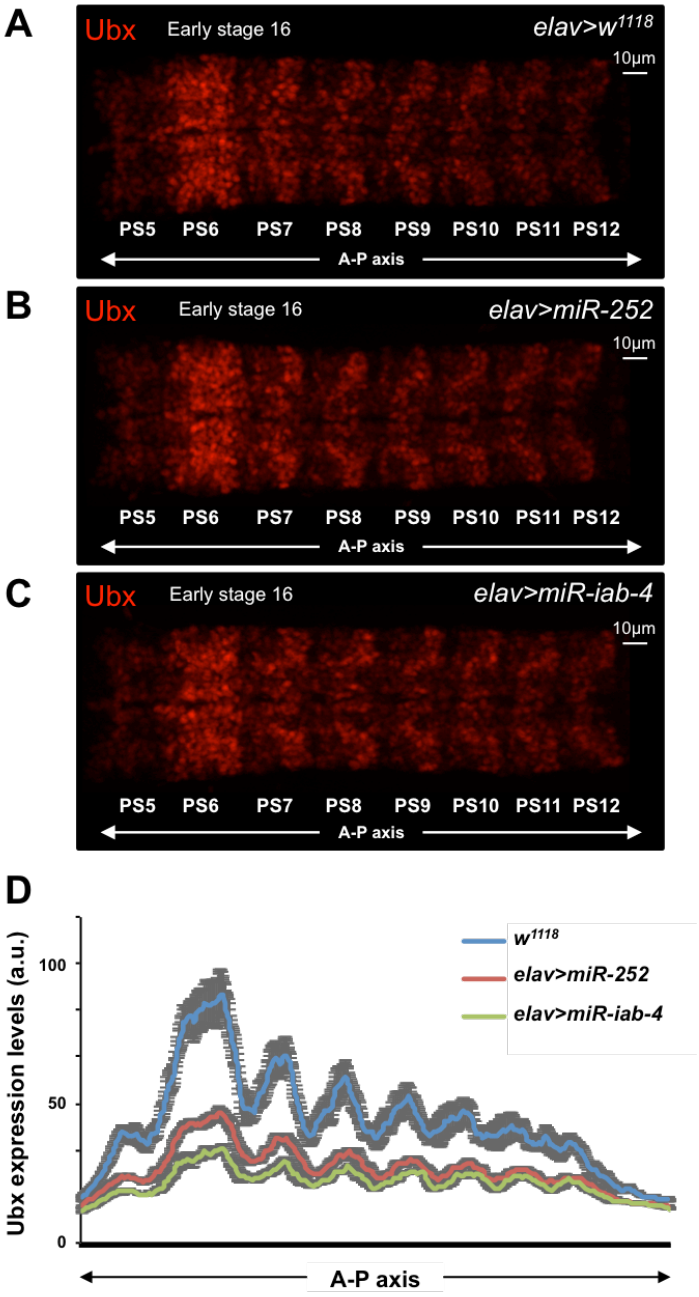


Figure 5.5 Ubx expression changes by ectopic expression of *miR-252* in the *Drosophila* CNS.

Ubx immunohistochemistry in early stage 16 *Drosophila* CNS for negative control of Ubx endogenous expression levels, w1118; +; +/ Elav>Gal4 (A), ectopic expression of *miR-252*, w1118; +; UAS-miR-252/ Elav>Gal4 (B), and positive control, w1118; +; UAS-miR-iab-4/ Elav>Gal4 (C). The *Drosophila* CNS anterior is to the left. (D) Ubx expression level quantification for Figure 5.4A, B, C by using plot profiling in ImageJ. The blue curve indicates the expression of Ubx in negative control that shows the highest expression level. The red curve indicates the expression of Ubx in *miR-252* ectopic expression CNS that has lower expression level compared to negative control. The green curve indicates the expression of Ubx in the positive control that shows the lowest expression level. The gray shadow around the curve is the error bar (S.E.M.). Thus, ectopic expression of *miR-252* is sufficient to repress Ubx expression in the *Drosophila* CNS. Scale bar represents 10µm.

5.2.2.2 *miR-252* co-expresses with *Ubx* in the glial cell of the *Drosophila* CNS.

In the previous section, I confirmed that *miR-252* is sufficient to repress *Ubx* expression in the *Drosophila* CNS. If *miR-252* and *Ubx* are co-expressed in the same cell, the chance of regulation between them will increase dramatically. To identify the cell lineages in which *miR-252* regulates *Ubx* expression, I used a double fluorescent *in situ* hybridisation approach to locate the cells where *miR-252* and *Ubx* are co-transcribed. The probe for *Ubx* is specially designed to hybridise with sequences within the third intron of the *Ubx* gene (Figure 5.6A), which can detect nascent RNA transcription of *Ubx* and as such provide a glimpse of *Ubx* expression before post-transcriptional regulation occurs. This experimental design greatly increases the potential for identification of cell lineages where *miR-252* and *Ubx* interaction may happen.

Using the aforementioned approach in conjunction with the Nuclear Envelope Marker Lamin, I found that *miR-252* and *Ubx* are co-transcribed in several cells. Next, I aimed to identify specific CNS cells where *miR-252* and *Ubx* are co-transcribed. Testing several cell lineage markers in the CNS, such as engrailed, even-skipped, and *repo*, I found that the general expression pattern of *miR-252* was quite similar to that of Repo, a glial cell marker (Figure 5.6B). *miR-252* and *Ubx* were co-transcribed in the same repo marked cells in parasegment 6 (Figure 5.6B). Based on the glial cell lineage analysis data from Beckervordersandforth et al (Beckervordersandforth et al. 2008), I deducted the co-expression cell lineage from the position of this glial cell. I suggest that *miR-252* and *Ubx* are co-transcribed in the dorsal lateral subperineurial glia (DL-SPG) that derives from abdominal neuroblast 5-6 lineage (NB5-6A) (Figure 5.6C, D).

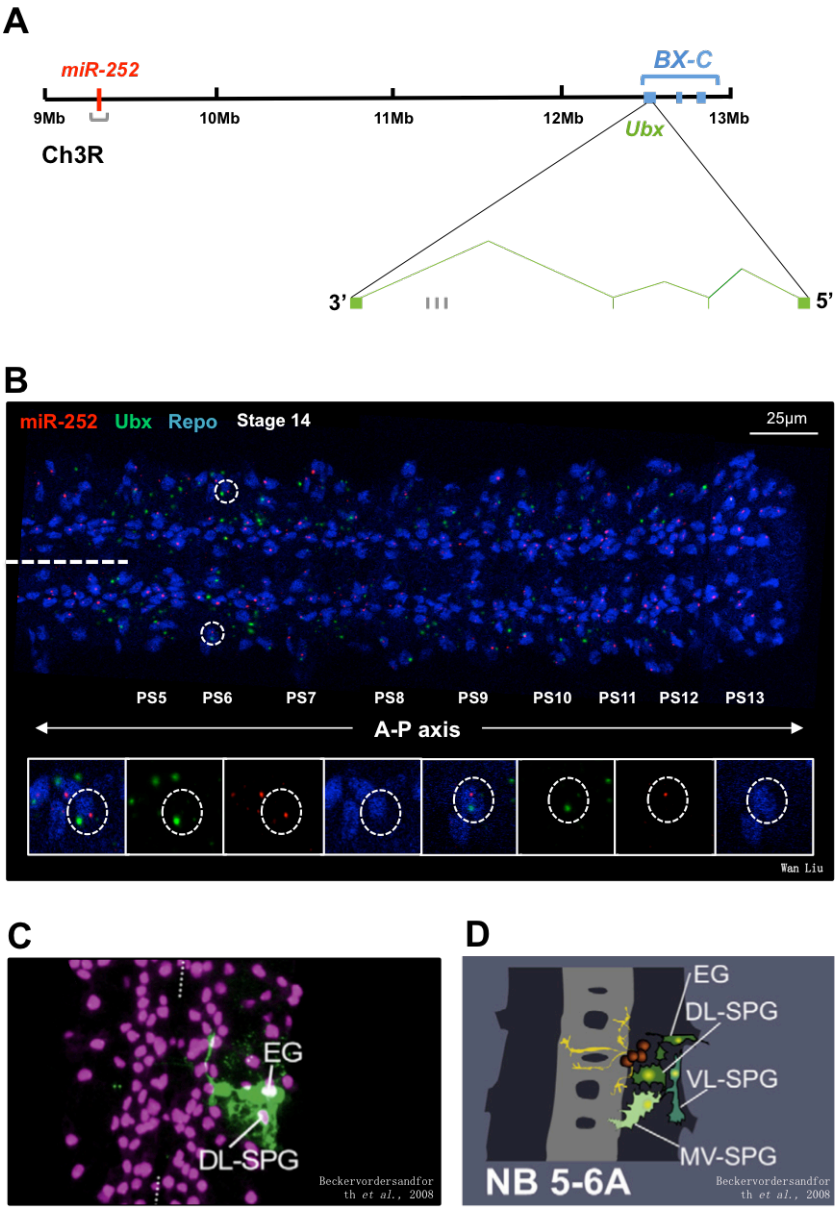


Figure 5.6 Co-expression of *miR-252* and *Ubx* in the *Drosophila* CNS.

(A) The diagram of the locus of *miR-252* and *Ubx* in the third chromsome righ arm denoting the probes designed for *miR-252* and *Ubx*. The probe of *miR-252* is the same as the one described in Figure 5.4A. The probe for detecting nascent *Ubx* transcripts was designed in the third intron of *Ubx* as described in the study of Rogulja-Ortmann *et al.* (Rogulja-Ortmann et al. 2014). (B) Double Fluorescent *in situ* hybridisations for *miR-252* and *Ubx* in stage 14 *Drosophila* CNS showing a single stack. The red dots show the expression pattern of *miR-252*. Green dots show nascent *Ubx* expression. The blue dots show Repo staining which marks glial cell nuclei. *miR-252* and nascent *Ubx* transcripts are co-expressed in the anterior of parasegment 6 (white circles). The *Drosophila* CNS anterior is to the left. (C) Dorsal stack of NB5-6A lineage shows DiI labelled exit glia (EG) and DL-SPG. Fluorescent DiI labelled clones in green, glial cells in magenta. (D) Cartoon presents the whole NB5-6A-derived clone, comprising neurons (cell bodies in red, axonal projections in yellow) in addition to glial cells (in green). Both image (C) and (D) are taken from Beckervordersandforth *et al*. (Beckervordersandforth et al. 2008). *miR-252* and *Ubx* appear to be co-transcribed in DL-SPG cells that are derived from NB5-6A. Scale bar represents 25μm.

5.2.2.3 *miR-252* represses Ubx expression in the glial cell of the *Drosophila* CNS.

From previous experiments, I demonstrated that *miR-252* is sufficient to repress *Ubx* expression in the *Drosophila* CNS (Figure 5.5). I now asked whether *miR-252* is necessary to repress *Ubx* expression in the *Drosophila* CNS. To answer this question, I need to test for *Ubx* expression level changes in the *Drosophila* CNS in a *miR-252* mutant background and compared it to a wild-type control. If *Ubx* expression levels increase in a particular cell of a *miR-252* mutant background *Drosophila* CNS where *miR-252* and *Ubx* are co-expressed (Figure 5.6B), I can conclude that *miR-252* is necessary to repress *Ubx* expression in the wild-type *Drosophila* CNS.

I used a fly stock from *Drosophila* Genetic Resource Center to answer this question. This fly stock has a *piggyback* element insertion in the 3' end of the *miR-252* hairpin, shown in Figure 5.7A. This large sequence insertion in the 3' end of the miRNA hairpin could affect the formation of the secondary structure and miRNA maturation. Zeng and Cullen reported that flanking nonstructured RNA sequence around primary transcription is required for Drosha processing of miRNA maturation (Zeng & Cullen 2005). In addition, it has been shown that a short stem of 2-3 nt 3' overhangs is necessary to be recognized by Exportin 5 during the process of pre-miRNA transport from nucleus to cytoplasm (Wahid et al. 2010). This large *piggyback* insertion in the 3' end of the *miR-252* hairpin will interfere with the proper formation of 3' overhangs. Thus, I hypothesized that the pre-miR-252 cannot be processed into a mature miRNA in the context of this *piggyback* insertion. I use an experimental approach to test this hypothesis, and thus the miRNA mutant status of the aforementioned stock in the following section.

The particular design of TaqMan RT-PCR, which consists of stem-loop RT and TaqMan PCR analysis, leads to a highly sensitive, specific and precise method to quantify mature miRNAs (Chen et al. 2005). Using the TaqMan RT-PCR technique, I quantified the expression levels of *miR-252-5p* in the *miR-252* mutant fly (Figure 5.7B). Additionally, I used the flies with *miR-252* ectopically

expressed by an *elav*-Gal4 driver, with wild-type flies as controls. It has been shown that *miR-252* is expressed during *Drosophila* development (Ruby et al. 2007). Theoretically, the *miR-252* expression levels should be increased in flies where *miR-252* is ectopically expressed, when compared to wild-type flies, while *miR-252* expression should be undetectable or reduced in *miR-252* mutant flies. Indeed, I found that the *miR-252-5p* expression level in *miR-252* mutant stock was almost non-existent, while the expression level of miR-252-5p showed a dramatic increase in flies where *miR-252* had been ectopically expressed when compared to that of the wild-type flies. This result validates that the point mutation is a *miR-252* mutant allele.

To have a brief idea of *miR-252* function during *Drosophila* development, I assessed the survival of *miR-252* mutant flies (Figure 5.7C). I found that *miR-252* mutant flies survive as well as wild-type flies throughout embryogenesis, first instar larva (L1), second instar larva (L2), and the third instar larva (L3) stages. However, there is almost 25% reduction in the number of *miR-252* mutant L3 larvae reaching the pupation stage, when compared to that of the wild-type, and the amount of eclosed flies in our *miR-252* mutant stock is about half of that of wild-type flies (Figure 5.7C). Therefore, I conclude that a loss of *miR-252* expression leads to reduced survival of the fly and that this miRNA does, as such, affect *Drosophila* development. More experiments have to be done to prove it.

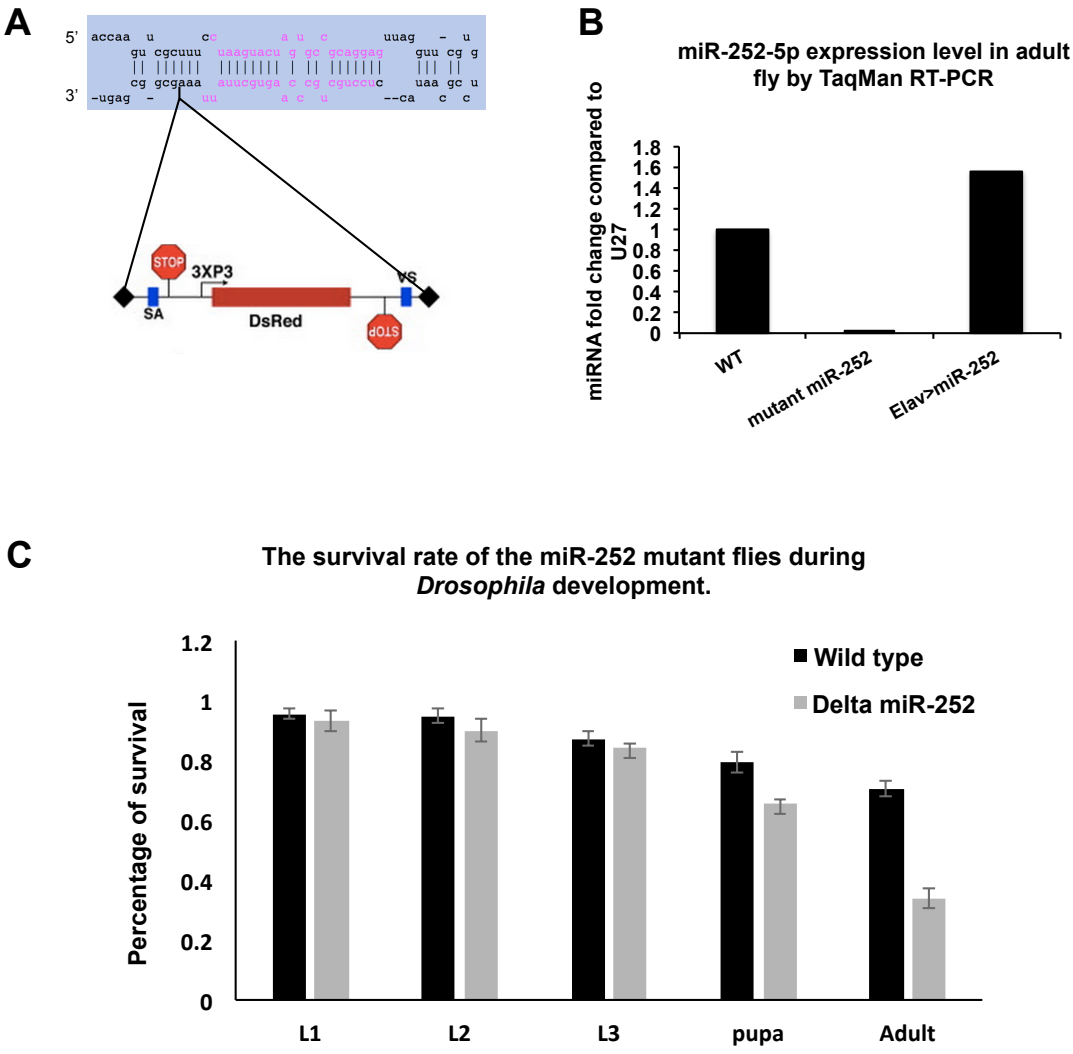


Figure 5.7 The *miR-252* mutant fly stock.

(A) *piggyback* element inserts into the 3' end of the *miR-252* precursor hairpin, which causes the *miR-252* mutation. The *miR-252* hairpin structure was modified from miRbase and the diagram depicting a *piggyback* mutator element was adapted from Schuldiner *et al.* (Schuldiner et al. 2008). (B) I used TaqMan RT-PCR to measure *miR-252* expression levels in wild-type flies, *miR-252* mutant flies and *elav*-Gal4>UAS-miR-252 flies. miRNA expression levels are normalized to U27, an artificial TaqMan miRNA control assay. (C) The survival rate of *miR-252* mutant flies during *Drosophila* development. The total number of embryos selected for the survival assay was 50 for each genotype without bias, and I calculated survival rates throughout *Drosophila* development, including first instar larvae (L1), second instar larvae (L2), third instar larvae (L3), pupal stages and adulthood. Wild-type flies were used as a positive control. Error bars represent the S.E.M. Loss of *miR-252* expression reduces fruitfly survival from pupation onwards, and hampers as such, fly development.

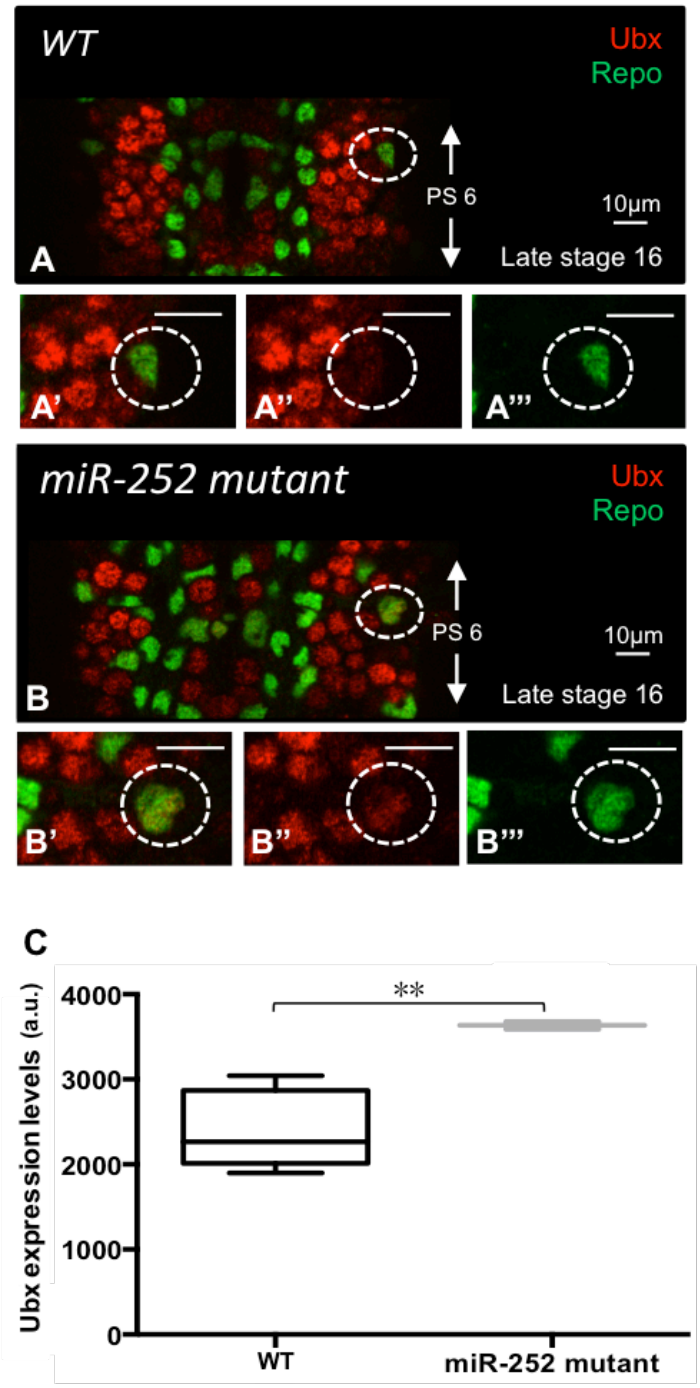


Figure 5.8 Analysis of Ubx expression levels in *miR-252* mutant *Drosophila* CNS.

Staining of Ubx and Repo (glial cell marker) in late stage 16 *Drosophila* CNS parasegment 6 and assessment of Ubx expression level changes between (A) wild-type and (B) *miR-252* mutant. Detailed analysis of higher maginification images (twofold) of DL-SPG in (A'') Ubx staining and (A''') Repo staining for wild-type. Detailed analysis of higher maginification images (twofold) of DL-SPG in (B'') Ubx staining and (B''') Repo staining for *miR-252* mutants. The *Drosophila* CNS anterior is to the top. (C) Quantification for the cell marked by the white dashed circle (n=5 per genotype). An Unpaired t-test, parametric with welch correction was performed to compare different genotype; $p<0.01$ (**). Error bars represent the error (S.E.M.). I found that *miR-252* is necessary to repress *Ubx* expression in the DL-SPG of the *Drosophila* CNS. Scale bar represents 10µm.

After I confirmed that our stock was indeed a *miR-252* mutant, I tested if *miR-252* controls *Ubx* expression in the *Drosophila* CNS. I quantified the Ubx expression levels in the glial cell lineage, DL-SPG where *miR-252* and *Ubx* are co-transcribed this *miR-252* mutant and wild-type (Figure 5.8A'-B'). The Ubx expression level was increased in *miR-252* mutant CNS, when compared to that of the wild-type (Figure 5.8C), which indicates that *miR-252* is necessary to repress *Ubx* expression in the DL-SPG of the *Drosophila* CNS.

5.2.3 The role of *miR-92a* in the *Drosophila* CNS.

Similarly to *miR-252*, *miR-92a* is also sufficient to repress *Ubx* expression in *Drosophila* haltere imaginal discs (Figure 4.8D', 4.9). Additionally, it is expressed in the *Drosophila* CNS, a tissue context where Ubx is present (Figure 5.2B). In order to examine the *miR-92a* expression pattern within the *Drosophila* CNS and test for a potential molecular interaction between *miR-92a* and *Ubx*, I combined fluorescent *in situ* hybridisations for *miR-92a* RNAs and antibody stainings to detect Ubx protein expression. *miR-92a* is expressed in the *Drosophila* CNS in a complementary manner to that of Ubx (Figure 5.9B). The cells where *miR-92* is expressed do not show Ubx protein expression. Based on this set of results, I suggest that there might be a *miR-92a* suppression effect on Ubx in the CNS.

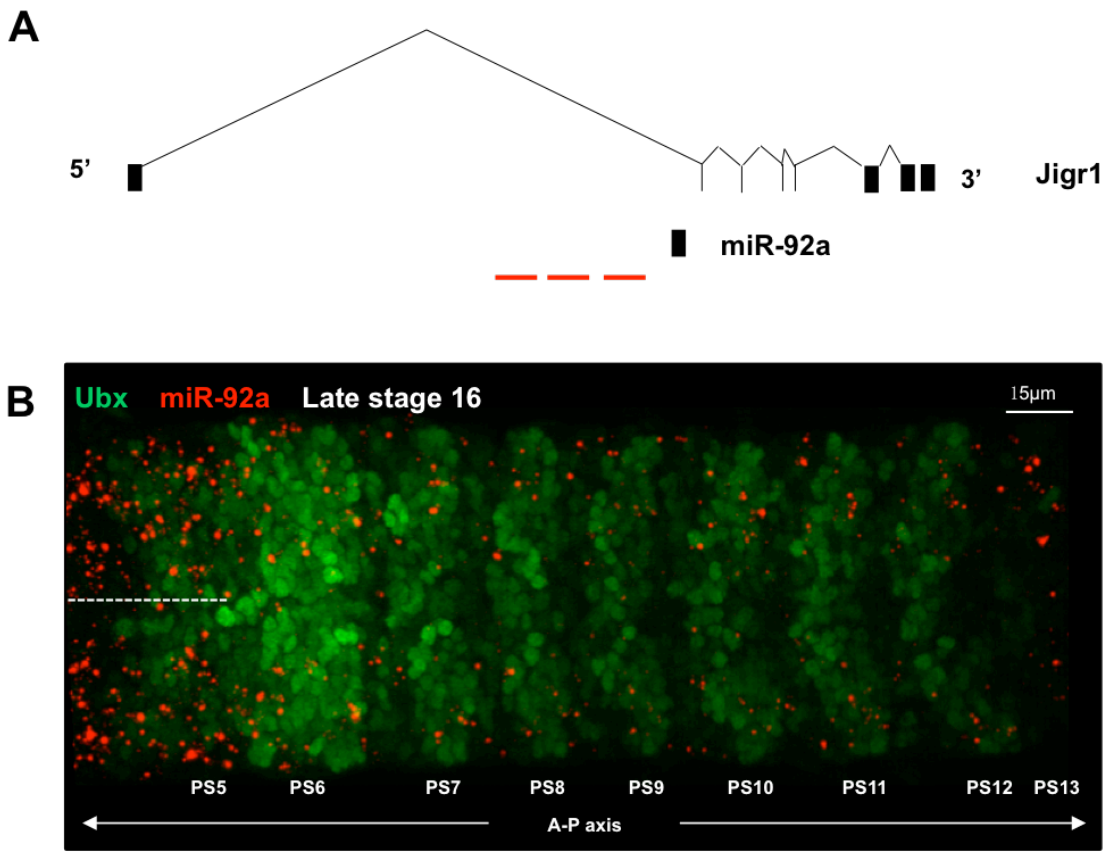


Figure 5.9 Spatial expression of *miR-92a*.

(A) Diagram of the locus of *Jigr1* denoting the probe designed for *miR-92a*. The filled black squares and the vertical lines are the exons for *miR-92a* host gene *Jigr1* and the single lines that connect the squares represents excised introns. The black square beneath *Jigr1* represents the position of the *miR-92a* primary transcription in the intron region of *Jigr1*. *miR-92a* and its host gene *Jigr1* are transcribed from left to right. The red lines beneath the sequence of the first intron region of *Jigr1* represent the annealing sites for the multiple probes designed to detect *miR-92a*. (B) Spatial expression of *miR-92a* in late stage 16 *Drosophila* CNS on the area where *Ubx* is also expressed. *miR-92a* is expressed in the whole CNS, but mainly in the brain (Figure 5.2B). The red dots show the expression pattern of *miR-92a*. Green dots represent Ubx expression in the CNS. The *Drosophila* CNS anterior is to the left. *miR-92a* and Ubx proteins are expressed in a complementary manner. Scale bar represents 15μm.

5.2.3.1 *miR-92a* is sufficient to repress Ubx expression in the *Drosophila* CNS.

miRNAs repress gene expression at the post-transcriptional level. If *miR-92a* regulates *Ubx* expression in the *Drosophila* CNS, it should display genetic sufficiency to repress *Ubx* expression. In order to test for an inhibitory role of *miR-92a* on Ubx expression, I used *elav*-Gal4 to overexpress *miR-92a* in the

Drosophila CNS and I measured Ubx expression levels (Figure 5.10B). Ubx expression levels, in the context of ectopic *miR-92a* expression in the CNS, are reduced when compared to the negative control (Figure 5.10A), confirming that *miR-92a* is sufficient to repress Ubx expression in the *Drosophila* CNS.

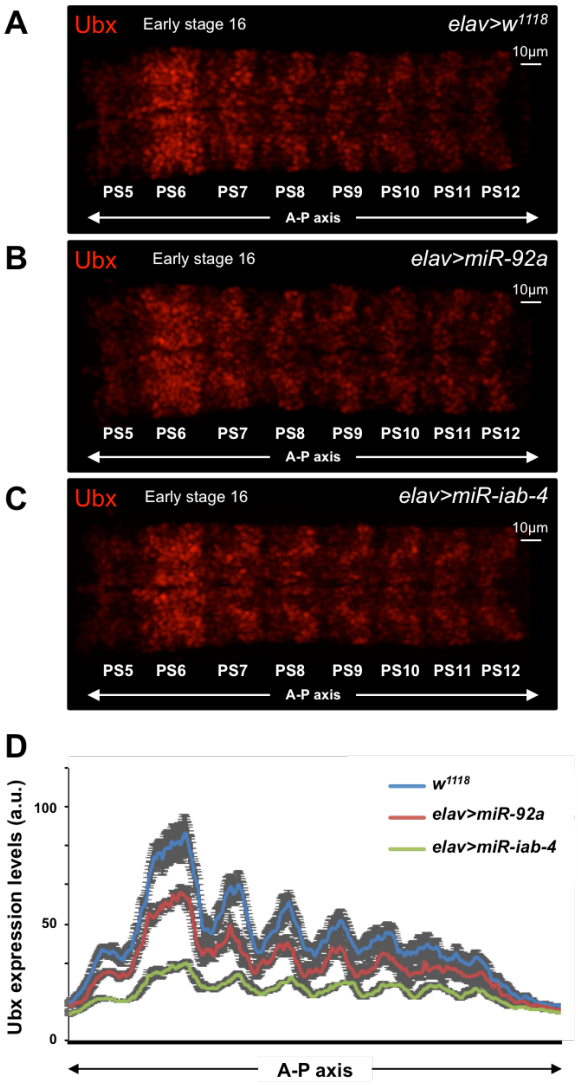


Figure 5.10 Ubx expression changes by ectopic expression of *miR-92a* in the *Drosophila* CNS.

Ubx immunohistochemistry in early stage 16 *Drosophila* CNS for negative control of Ubx endogenous expression levels, w1118; +; +/ Elav>Gal4 (A), ectopic expression of *miR-92a*, w1118; +; UAS-miR-92a/ Elav>Gal4 (B), and positive control, w1118; +; UAS-miR-iab-4/ Elav>Gal4 (C). The *Drosophila* CNS anterior is to the left. (D) Ubx expression level quantification for Figure 5.10A, B, C by using the plot profile function in ImageJ. The blue curve indicates Ubx expression in the negative control, which shows the highest expression level of all treatments. The red curve represents the expression of Ubx in CNS with ectopically expressed *miR-92a*, which has lower expression level compared to negative control. The green curve indicates the expression of Ubx in the positive control that shows the lowest expression level. The gray shadow around the curve is the error bar (S.E.M.). Thus, ectopic expression of *miR-92a* is sufficient to repress Ubx expression in the *Drosophila* CNS. Scale bar represents 10μm.

5.2.3.2 *miR-92a* is not necessary to repress Ubx expression in the engrailed marked cell lineage of the *Drosophila* CNS.

In the previous section, I confirm that *miR-92a* is sufficient for the repression of *Ubx* expression in the *Drosophila* CNS. Next, I wanted to examine whether *miR-92a* is necessary to repress *Ubx* expression in the *Drosophila* CNS. If the answer to this biological question is affirmative, I could conclude that *miR-92a* has a regulatory role on *Ubx* in the *Drosophila* CNS. In order to understand if *miR-92a* is necessary to repress *Ubx* expression in the *Drosophila* CNS, I assessed Ubx expression levels in the *miR-92a* deficiency stock (Figure 5.11B). I found that the Ubx expression levels were similar between the CNS of *miR-92a* deficiency mutants and wild-type controls (Figure 5.11). Thus, it is possible that *miR-92a* does not regulate *Ubx* expression in the *Drosophila* CNS.

However, it is also possible that *miR-92a* is necessary for the appropriate regulation of *Ubx* expression in the *Drosophila* CNS, and that this is hindered by our data collection approach. I have observed that *miR-92a* is expressed in a limited number of cells in the *Drosophila* CNS (Figure 5.9B), and it is possible that there is a change of Ubx expression level in these particular CNS cells when the wild-type is compared to the *miR-92a* mutants. These minute differences could be undetectable to us due to the variation of the Ubx expression from the large amount of other Ubx-expressing CNS cells. Therefore, I decided to further identify the cells where *miR-92a* and *Ubx* are co-transcribed in order to examine Ubx protein expression level changes in that particular cell lineage.

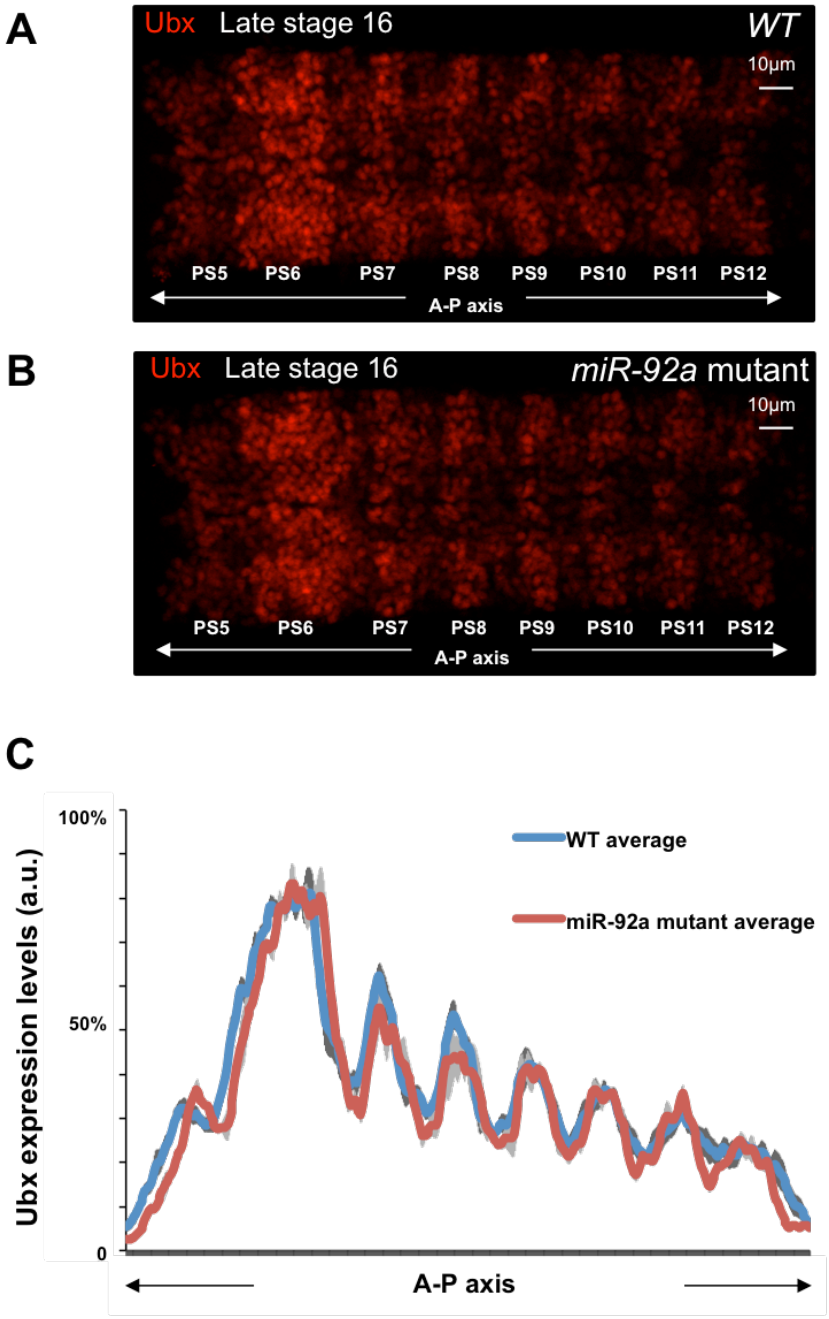


Figure 5.11 Analysis of Ubx expression levels in *miR-92a* mutant *Drosophila* CNS.

Antibody staining for Ubx in late stage 16 *Drosophila* CNS, comparing Ubx expression levels between (A) wild-type and (B) *miR-92a* mutant. (C) Quantification of the Ubx expression levels in Figure 5.11A, B. The red curve depicts the expression levels of Ubx in *miR-92a* mutants while the blue curve displays the expression levels of Ubx in wild-type flies. The light and dark gray shadows represent the error (S.E.M.) for both *miR-92a* mutants and wild-type, respectively. The *Drosophila* CNS anterior is to the left. No Ubx expression level changes in the *Drosophila* CNS are observed, when the wild-type and *miR-92a* deficient flies are compared. Scale bar represents 10µm.

After trying several cell-lineage markers in the *Drosophila* CNS, I found that *miR-92a* was co-expressed with Engrailed in one particular cell (probably derived from NB6-2) (Doe 1992) of parasegment 6 (Figure 5.12). It has been reported that *Ubx* expression in the *Drosophila* embryo partially overlaps with that of *engrailed* (Mann 1994; Qian et al. 1993). So I re-analysed Ubx expression level changes in a *miR-92a* mutant background, this time focusing in this particular cell (Figure 5.13). The Engrailed expression pattern in parasegment 6 has been shown to vary across individuals (Figure 5.12A', B', C'); I identified three main modes of Engrailed expression patterns. Using this approach, I identified cells where *miR-92a* and Engrailed are co-expressed (Figure 5.12D, E, F). When I analysed Ubx expression level change between wild-type CNS and *miR-92a* mutant CNS, I considered all these three situations together. However, I did not find that Ubx expression levels change between them (Figure 5.13A''', B''', C). Therefore, *miR-92a* appears not to be necessary to repress *Ubx* expression in an Engrailed-marked cell lineage in the *Drosophila* CNS.

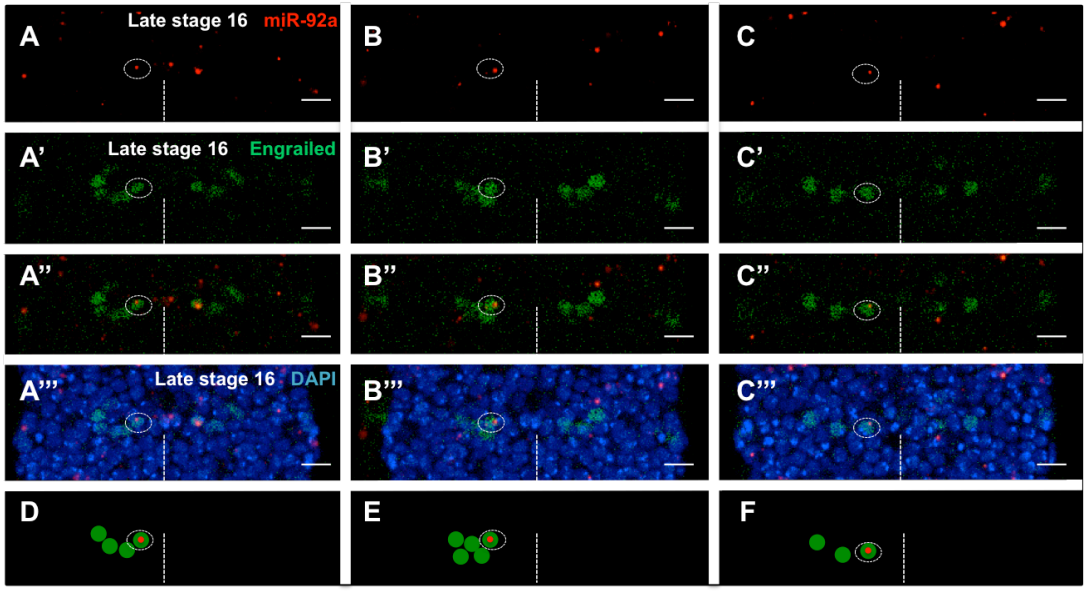


Figure 5.12 Identification of the neural lineages where *miR-92a* is expressed.

Fluorescent *in situ* hybridisation of *miR-92a* and the subsequent antibody staining for Engrailed in parasegment 6 of late stage 16 *Drosophila* CNS. (A, B, C) Fluorescent *in situ* hybridisation for *miR-92a*. (A', B', C') Engrailed staining in the *Drosophila* CNS. (A'', B'', C'') Merged images of A, B, C and A', B', C'. (A''', B''', C''') A'', B'', C'' with DAPI staining. (D, E, F) Three different Engrailed expression patterns in parasegment 6 of late stage 16 *Drosophila* CNS are represented in green, and the red dots identify the cells where *miR-92a* co-transcribes with *engrailed*. These exist three different modes of *miR-92a* co-expression with Engrailed. *Drosophila* CNS anterior is to the top. Scale bar represents 10μm.

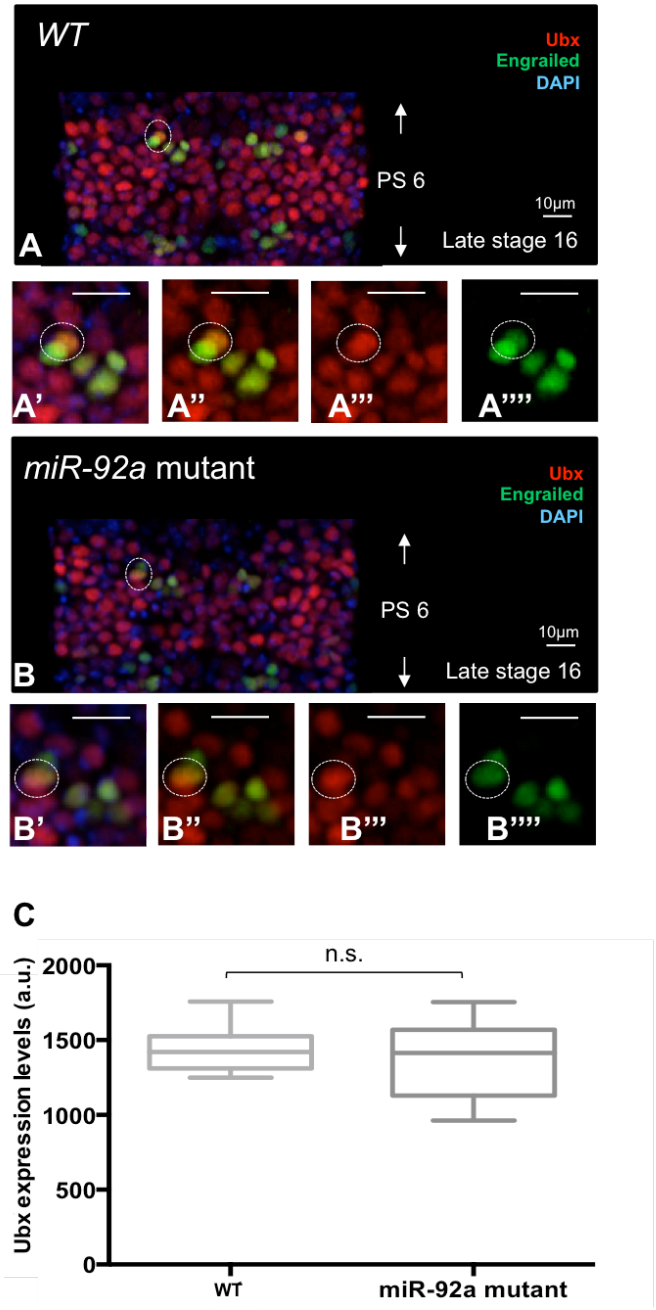


Figure 5.13 Analysis of Ubx expression levels in Engrailed positive miR-92a-expressing cell lineage within the *Drosophila* CNS.

Staining of Ubx, Engrailed (cell marker) and DAPI in late stage 16 *Drosophila* CNS parasegment 6 and compared the Ubx expression level changes between (A) wild-type and (B) *miR-92a* mutant. Detailed analysis of higher maginification images (twofold) of NB6-2 derived Engrailed marked cell in (A''') Ubx staining and (A'''') Engrailed staining for wild-type. Detailed analysis of higher maginification images (twofold) of NB6-2 derived Engrailed-marked cell in (B''') Ubx staining and (B'''') Engrailed staining for *miR-92a* mutants. The *Drosophila* CNS anterior is to the top. (C) Quantification for the cell marked by the white dashed circle (n=5 per genotype). An Unpaired *t-test*, parametric with welch correction was performed to compare the results for different genotypes; $p>0.05$ (n.s.). Error bars represent the error (S.E.M.). I found that *miR-92a* is not necessary to repress *Ubx* expression in the Engrailed-marked cell derived from NB6-2 of the *Drosophila* CNS. Scale bar represents 10µm.

5.3 Discussion.

In this chapter, I investigated the effects of miRNA regulation on the expression of *Hox* gene *Ubx* within the *Drosophila* CNS. For this purpose, I developed a cellular approach to study the effects of miRNA expression on the activity of *Hox* genes within the CNS. Most of this work is focused on the *Drosophila Hox* gene *Ubx*, as this is one of the *Hox* genes for which post-transcriptional regulation is currently best understood.

To this end, I asked the following biological question: what candidate miRNAs might regulate *Hox* gene *Ubx* expression during embryonic neural development? I examined both the temporal and spatial expression pattern of the miRNAs that previously identified as potential regulators of *Ubx* in the haltere.

First, I inspected the miRNA RNA-seq data from Ruby *et al.* (Ruby et al. 2007). To have an overview of developmental temporal expression patterns for 9 *Drosophila melanogaster* miRNAs, *miR-92a*, *miR-92b*, *miR-303*, *miR-304*, *miR-252*, *miR-iab-4*, *miR-310*, *miR-311*, and *miR-312*. I found that *miR-303* and *miR-304* were lowly expressed during embryogenesis and *miR-iab-4* has previously been shown to regulate *Ubx* expression in the *Drosophila* CNS (Thomsen et al. 2010; Bender 2008). Therefore, I excluded these three miRNA for further analyss.

Next, I used a combination of *in situ* hybridisation and immunohistochemistry techniques to measure the temporal expression patterns for the other 6 miRNAs, and found that only *miR-252* and *miR-92a* were expressed in a complementary manner with *Ubx* in the *Drosophila* CNS during embryogenesis. Thus, these miRNA species are great candidates for further analysis. *miR-310C* is expressed in complementary manner with Ubx in the larval CNS, suggesting that this miRNA species is a good candidate to study the mechanism of molecular control of *Ubx* activity by miRNAs in the larval CNS. This aspect of the study is, however, reserved for future work.

After showing that both *miR-252* and *miR-92a* are sufficient to repress Ubx expression in *Drosophla* CNS, I addressed the following two questions: (i) How do *Ubx* expression levels relate to those of candidate miRNAs at the level of single cells? (ii) What are the specific neuronal lineages showing changes of *Ubx* expression when miRNAs are mutated?

For the case of *miR-252*, I identified that *miR-252* and *Ubx* are co-expressed in DL-SPG cells, which derive from NB5-6A. In analysis of the *miR-252* mutants, I found that Ubx expression levels are increased in glial cells. Thus, *miR-252* is sufficient and necessary to repress *Ubx* expression in DL-SPG cell of the *Drosophila* CNS.

From our previous experiments, I found that there is an indirect interaction between *miR-252* and *Ubx* in *Drosophila* haltere imaginal discs (Figure 4.15G', I). Does *miR-252* also regulate *Ubx* indirectly in the *Drosophila* CNS? I can apply the same strategy that I used in the haltere imaginal discs, but instead of using *sd*-Gal4, I would use *repo*-Gal4 to ectopically express a UAS-mCherry.*Ubx*.long.3' UTR and UAS-*miR-252* in the glial cells of the *Drosophila* CNS. In this experimental scenario, mCherry expression could be measured in the DL-SPG where *miR-252* is ectopically expressed and compared to a control without a *miR-252* overexpression construct. If *miR-252* directly targets *Ubx*.long.3' UTR, a reduction of mCherry expression should be observed. If *miR-252* indirectly regulates *Ubx* expression in DL-SPG of the *Drosophila* CNS, I suggest that *miR-252* regulates other *Ubx* related genes to control *Ubx* expression.

In the case of *miR-92a*, I identified that *miR-92a* is expressed in the Engrailed-marked cell that derives from NB6-2. I did not find any Ubx expression-level reduction in this cell in *miR-92a* mutant CNS when compared to the wild-type control. Thus, *miR-92a* is not necessary to regulate *Ubx* expression in this Engrailed-positive cells.

However, I have considered the following caveats that may mask the *Ubx* expression level change between wild-type and *miR-92a* mutant CNS. First, I

identified the cell lineage relying solely on the relative position of the cell when compared to all Engrailed-positive cells. Without more experimental work to further investigate the specific cell lineage that contains this cell of interest, I cannot rule out that the same cells expressing *miR-92a* are ones used for the quantification of Ubx protein expression levels. Secondly, I studied this cell in three variable modes of expression-patterns for Engrailed, which may include several different cells at the same time. Thirdly, the Ubx has negative autoregulation (Crickmore et al. 2009). The lack of *miR-92a* causes high Ubx expression that is repressed by Ubx itself in the *miR-92a* mutant fly. Thus, I could not observe the Ubx expression increase in the *miR-92a* mutant CNS. Lastly, this is only one experimental result with limited samples. More biological repeats need to be performed in order to draw a more solid conclusion.

In summary, I found that *miR-252* is sufficient and necessary to repress *Ubx* expression in the DL-SPG cells of the *Drosophila* CNS, which opens up the possibility to address the question of what is the biological role of miRNA-mediated post-transcriptional regulation during CNS development.

CHAPTER 6 General discussion

6.1 Introduction

Mutations in the proteins required for miRNA biogenesis or function have been shown to affect animal development (Okamura et al. 2004; Liu et al. 2004; Wienholds et al. 2003; Grishok et al. 2001; Lee, Y. S.et al. 2004; Bernstein et al. 2003). Specifically, several miRNAs play essential roles in cellular and developmental processes, such as developmental timing, cell death and proliferation, morphogenesis and cell differentiation (Ambros 2004; Alvarez-Garcia & Miska 2005). Therefore, miRNA-mediated post-transcriptional regulation is a crucial layer of gene regulation during animal development. However, the mechanisms that allow miRNAs to identify their targets in the context of development are still unclear.

The *Hox* gene *Ubx* encodes a homeodomain transcription factor that controls the genetic networks underlying haltere development in *Drosophila*. Among the known 256 *Drosophila* miRNAs, only *miR-iab-4* and *miR-iab-8* are proven to regulate *Ubx* expression during *Drosophila* development (Bender 2008; Ronshaugen et al. 2005; Stark et al. 2008; Thomsen et al. 2010; Tyler et al. 2008). The relationships between *Ubx* and the other 254 miRNAs are not clear. Are there any other *Hox* regulatory miRNAs in *Drosophila*? Are there any fast and easy approaches to identify *Hox* regulatory miRNAs? Given that the reduction of *Ubx* expression causes easily trackable morphological changes in the adult haltere, I decided to carry out a genetic screen to assess the potential effects of miRNA ectopic expression on *Ubx* repression, using the analysis of haltere phenotypes as a proxy for changes in *Ubx* expression.

6.2 *Hox* cluster miRNAs

In total, I screened 96 miRNAs and 10 miRNA clusters transgenes. Including *Hox* cluster miRNAs (*miR-iab-4* and *miR-10*), 28% of miRNAs (27 miRNAs) and 10% of miRNA clusters (1 miRNA clusters) screened have the potential to

regulate *Hox* gene *Ubx* expression in the haltere.

Most evidence of miRNA-mediated *Hox* gene regulation comes from the miRNAs encoded within the *Hox* cluster itself, which has been shown to regulate *Hox* gene expression during animal development. In vertebrates, only *miR-10* and *miR-196* have been experimentally proven to regulate *Hox* gene expression during development (Yekta et al. 2004; He et al. 2012; Woltering & Durston 2008; McGlinn et al. 2009). For example, *miR-196* directs cleavage of *HOXB8*, leading to the repression of this *Hox* gene expression in developing mouse embryos (Yekta et al. 2004). In *Drosophila*, *miR-iab-4* and *miR-iab-8* regulate *Ubx* expression (Bender 2008; Garaulet et al. 2014; Gummalla et al. 2014; Ronshaugen et al. 2005; Stark et al. 2008; Thomsen et al. 2010; Tyler et al. 2008).

In this study, I showed that six miRNAs directly target *Ubx* 3' UTR, five of which are not encoded within *Hox* clusters. Thus, *Hox* genes are not only specifically regulated by *Hox* cluster miRNAs. Conversely, *Hox* cluster miRNAs do not exclusively regulate *Hox* genes. For example, *miR-993* is encoded in *Hox* cluster, but ectopic expression of this miRNA shows no haltere phenotype (Figure 3.7N). In addition, some of the screened miRNAs are evolutionarily conserved between vertebrate and invertebrate in the similar tissue. For example, *miR-92a* is expressed in CNS in both *Drosophila* (Figure 5.2B) and mice (Saugstad 2010). The validated miRNA-mRNA interaction in *Drosophila* could be applied to the study of that in the vertebrate as well.

6.3 *miR-92* family

Interestingly, the haltere imaginal disc overexpression of miRNAs in the *miR-92* family leads to loss of the cuticle trichomes that cover the haltere surface. This feature is consistent with previous reports that ectopic expression of *miR-92a* and its seed relatives have the unique ability to trigger trichome loss (Bejarano et al. 2012; Schertel et al. 2012). However, miRNAs from *miR-92* family show different levels of homeotic transformation, which suggests that besides seed

sequence, other factors may affect the targeting of miRNAs in the *miR-92* family. It has been shown that miRNA 3' end determines target specificity within miRNA family (Brennecke et al. 2005). By using phylogenetic analysis and multi-species alignments of the mature miRNA sequences, I found that miRNAs that produce similar phenotypes tend to cluster together indicating that they might play similar roles on post-transcriptional regulation. The alignment of the mature miRNA sequence shows that the t1A position plays a very important role in miRNA-mediated repression, consistent with previous reports that nucleotides at the t1A position tend to be conserved in many mammalian miRNA target sites and enhance miRNA-mediated repression (Bartel 2009; Lewis et al. 2005). Due to the similarity on their seed sequences, miRNAs from the same family have similar targets and therefore functions. However, this picture is blurred by the target affinity for miRNAs in the same family, which may vary.

6.4 Genuine miRNA target sites.

Compared to experimental approaches, bioinformatic predictions of miRNA targeting are a fast and flexible way to find out miRNA-target interactions. There are at least ten bioinformatics programs that can be used for miRNA prediction in *Drosophila*. To what extent can the current bioinformatic prediction tools help us in finding the genuine miRNA target sites on the 3' UTR?

To answer this question, I first uncovered a set of miRNAs that directly interact with *Ubx* 3' UTRs, and then I searched for bioinformatic predictions to validate them. Consistently with previous reports, I found that there is a major problem with the bioinformatics programs that is their high false positive rate (Witkos et al. 2011). Although the addition or more prediction parameters can increase the specificity of miRNA prediction algorithms, their sensitivity is also expected to concomitantly decrease (see detail analysis in 6.6).

However, a combination of experimental approaches and bioinformatics predictions could be a good way to finding genuine target sites for miRNAs on

3’ UTR *in vivo* and *in vitro* by: (i) ectopically expressing miRNAs in a developing organ (ii) screening for phenotype of a particular gene loss-of-function; (iii) using immunohistochemistry and genetic constructs to directly validate miRNA-mRNA targeting; (iv) comparing the results to bioinformatic predictions in order to generate precise hypotheses about specific miRNA target sites; (v) using genetic constructs where a reporter ORF is fused to a 3’ UTR with a mutated miRNA target site, in a cell-culture (e.g. S2 cells) or developmental (transgenic flies) to validate the result. By repeating the above steps, I could potentially catalogue all miRNA target sites on the *Ubx* 3’ UTR for all six miRNAs that directly target *Ubx*. Further experiment as described in step (v) still need to be carried out in the future, to validate this approach.

Genetic screen has been approved to be a very fleasible tool to study the miRNA function. The studies of miRNA functions from other groups have screened 44% (Bejarano et al. 2012), 43% (Schertel et al. 2012) and 50% (Szuplewski et al. 2012) miRNAs phenotypic positively from their genetic screen and proved miRNAs from the screened phenotypic positive miRNAs pools are related to the miRNA regulation of the studied genes. In addition, Ubx polyadenylation signals that control APA in cis are conserved across 12 different *Drosophila* species (Patraquim et al. 2011) and approximately a third of mouse Hox genes (many of which are evolutionarily conserved between mouse and human) produce a spectrum of mRNA isoforms generated by different types of RNA processing (Mallo & Alonso 2013), suggesting the aforementioned methodology for finding genuine target sites for miRNAs on 3’ UTR can be used to study other genes with clear phenotypic change when mutated in *Drosophila*, even in other systems and therefore it would aid to answer fundamental questions within developmental genetics.

6.5 *Ubx* regulatory miRNAs in the CNS

The miRNA milieu is unique for each cell type and has thus, in principle, the ability to shape the differential mRNA expression that is observed across cell types (Bartel & Chen 2004). Cell specific miRNA profiles have been shown to

refine the molecular responses that control cell differentiation and cell fate decisions on multiple tissues (Ivey & Srivastava 2010; Shenoy & Blelloch 2014). What is the mechanism of miRNA-mediated molecular control of *Hox* gene activity at the cellular level? Due to the very limited fate-mapping information regarding the developing haltere imaginal disc at cellular level (Roch & Akam 2000; Weatherbee et al. 1998), I tried to answer this question in the *Drosophila* CNS, where cell lineages to some extent are stereotypical and very well defined (Bossing et al. 1996; Schmidt et al. 1997; Schmid et al. 1999; Doe 1992; Beckervordersandforth et al. 2008; Ito et al. 1995; Manning et al. 2012).

Combining RNA-seq data for miRNA expression in *Drosophila* embryos (Ruby et al. 2007) and our *in situ* hybridisation results, I found that *miR-92a* and *miR-252* are potential regulators of *Ubx* expression in the *Drosophila* CNS. I showed that *miR-252* and *Ubx* are co-expressed in DL-SPG cells that derive from the NB5-6A lineage. By analysing *miR-252* mutants, I found that Ubx expression levels show an overall increase in the glial cells of *miR-252* mutant CNS. Thus, I conclude that *miR-252* is sufficient and necessary to repress *Ubx* expression in the DL-SPG cells of the *Drosophila* CNS.

It has been reported that glial cell are crucial regulators for nervous system development, function and health (Freeman & Doherty 2006). For example, glial cells provide the blood-brain barrier and allow the fine tuned homeostasis of ions and other small molecules (Stork et al. 2008). The discovery of *miR-252-Ubx* regulation in the DL-SPG glial cells might provide an insight into the miRNA effects on the glial cell biology.

6.6 Comparisons of the features of different bioinformatic prediction tools.

Understanding the mechanism of miRNA target site recognition remains a big challenge. One reason for this difficulty is the limited complementarity between miRNAs and mRNAs that mediates the interaction between a miRNA and its targets. Therefore, it is very likely that I can find hundreds of potential target sites per miRNA and a specific mRNA may also be predicted to be targeted by a large number of miRNAs. A number of bioinformatic and experimental

methods have been developed to study how miRNAs identify their targets. As I have shown in Chapter 3 and Chapter 4, each miRNA prediction program has its own specificity. Some use perfect base paring and evolutionary conservation as criteria, such as TargetScan and EMBL; some use a thermodynamic analysis, such as PITA and RNAhybrid; some consider the combined score of binding to different regions of the mRNA, such as DIANA; some combine different algorithms from different programs, such as miRanda and ComiR; some identify a specific 'target island' before identifying miRNA target sites; others were designed for combinatorial miRNA targeting prediction, eg. PicTar. Are any of the aforementioned parameters, claimed to be important for miRNAs target prediction, really important for miRNAs to find their targets? Which miRNA target prediction program is most reliable for prediction of *Drosophila* miRNA target sites?

Here, I have focused on analysing the miRNA targeting features of four bioinformatic programs that use the same miRNA input and the same 3' UTR isoform of *Ubx* in their predictions (Table 6.1). For these four programs, RNAhybrid shows the best performance, predicting targets on *Ubx* 3' UTRs for all 12 genuine miRNAs that experimentally proved to directly regulate *Ubx* expression (Figure 4.18). This program only uses the energetically most favorable hybridisation sites of a short RNA, combining sequence complemetarity and length of 3' UTR for the prediction. The limited parameters used in the prediction probably increase the sensitivity. However, it has the highest false positive rate among all four programs that were analysed.

The second-best program in our analyses is PITA, which predicts 9 out of 12 *Ubx* 3' UTR genuine target miRNAs. Compared to RNAhybrid, this program adds two more parameters into consideration for the prediction: $\Delta\Delta G$ and the energy costs for the unpairing of the flanking sequence. The extra parameters increase the specificity of PITA compared to that of RNAhybrid. However, the false positive rate of PITA is still higher than both miRanda and RNA22.

Table 6.1 Summary of the features used by the *Drosophila* target prediction programs considered in this study.

miRNA target prediction programs		miRanda	PITA	RNA22	RNAhybrid
Sequence	**Sequence complementarity**				
	5' dominant (canonical and seed)	Y	Y	Y	Y
	3'compensatory	Y	N	Y	N
	7-8 seed pairing with 1G:U, 1 bugle or 1 mismatch	Y	Y[e]	Y[f]	Y[g]
	Flanking				
	AU flanking	Y	N	N	N
	flanking sequence	N	Y[d]	N	N
	Target motif				
	Target motif	N	N	Y[a]	N
	Target position				
	relative distance of target position from end of 3' UTR	Y	N	N	N
	Cooperative regulation				
	cooperative regulation	Y	N	N	N
Thermodynamics	ΔΔG	N	Y	N	N
Conservation	target site conservation	Y[b]	N	N	N
Others	in vivo evidence	Y[c]	Y[c]	N	N
	length of 3' UTR	Y	N	Y	Y

The prediction parameters listed above are used for all miRNA prediction programs in this study. 'Y' indicates that the feature is used in the program, while 'N' means the feature is not considered. (a. use the miRNA features to identify the target islands in the mRNAs; The miRNA sequence features identify Target island, while miRNA target motif is pairing to the known miRNA sequences. b. the conservation of target sites is as a feature rather than a filter; c. these programs use *in vivo* experimental data as a validation rather than an approach to develop the program; d. considers the energy costs for the unpairing of the flanking sequence; e. only tolerance for G:U in the seed region; f. no restriction of G:U number in the seed region; g. only tolerate for one bugle in the seed region.)

miRanda detects 4 out 12 experimentally validated miRNA targeting events, while RNA22 recovers 3 out of the aforementioned 12 mature miRNAs, respectively. As such, the sensitivity and specificity of miRanda are both better than RNA22. Besides the common feature with RNA22, miRanda also takes AU flanking, relative distance of target position from end of 3' UTR, and cooperative regulation into account. It has been shown that AU sequences flanking the seed-complementary sequence are usually conserved in five vertebrate species (human, mouse, rat, dog and chicken), which indicates that this feature could be

important for miRNA target-site recognition (Lewis et al. 2005). Gaidatzis *et al.* report that miRNA target sites have the propensity to locate close to the start and end of 3' UTRs (Gaidatzis et al. 2007). Multiple RISCs in a physically close distance are hypothesized to cooperatively enhance miRNA regulation (Saito & Sætrom 2010). Therefore, adding these three features to the prediction would enhance the specificity. Also, miRanda uses *in vivo* evidence as a form of validation of the predictions, which is consistent the *in vivo* approach I used for the genetic screen. However, miRanda also takes target site conservation into consideration as a form of validation, an approach that may ignore all non-conserved target sites and thus decrease the sensitivity of the analysis.

In summary, the higher sensitivity of the miRNA prediction is always with the cost of the lower specificity. Some prediction factors, such as AU flanking, relative distance of target position from end of 3' UTR, and cooperative regulation, are valuable to take into account to improve the predicition sensitivity and specificity.

6.7 Concluding remarks

In this thesis, I described, first, a genome-wide gain-of-function screen of 96 miRNAs and 10 miRNA clusters, and found that 28% miRNAs and 10% miRNA clusters could induce *Ubx*-like mutant phenotypes in the halteres of adult *Drosophila melanogaster* flies. This study provides significant insights into potential miRNA-*Hox* regulatory interactions in a developmental context by manipulating miRNAs *in vivo*. Secondly, I employed a genetic approach to scan the physical interactions between the *Ubx* 3' UTRs and miRNAs and found six miRNAs in Class I directly interacting with *Ubx* 3' UTRs. Five of them are expressed in the haltere, indicating a potential biological role for these miRNAs in haltere development itself. Among them, only *miR-iab-4* is a *Hox* cluster miRNA. In addition, I used these *in vivo* experimental data to evaluate programs that are available for *Drosophila* miRNA target predictions, and search for genuine miRNA target sites in the *Ubx* 3' UTR. Lastly, I found that *miR-252* is sufficient and necessary to repress *Ubx* expression in the DL-SPG cells of the

Drosophila CNS, making it possible to address, in the future, the biological significance of miRNA-mediated post-transcriptional regulation during CNS development. My work thus contributes to the understanding of miRNA-mediated *Hox* gene regulation and, more generally, to the study of miRNA-target interactions within the physiological context of metazoan development.

BIBLIOGRAPHY

Abdelmohsen, K. et al., 2010. *miR-519* suppresses tumor growth by reducing HuR levels. *Cell Cycle*, 9, pp.1354–1359.

Akam, M., 1987. The molecular basis for metameric pattern in the *Drosophila* embryo. *Development*, 101(1), pp.1–22.

Akam, M.E. & Martinez-Arias, a, 1985. The distribution of Ultrabithorax transcripts in *Drosophila* embryos. *The EMBO journal*, 4(7), pp.1689–1700.

Alberts, B., Johnson, A. & J, L. et al., 2002. *Molecular Biology of the Cell* 4th ed., New York: Garland Science.

Alexiou, P. et al., 2009. miRGen 2.0: A database of microRNA genomic information and regulation. *Nucleic Acids Research*, 38, pp.137–141.

Alonso, C.R., 2012. A complex “mRNA degradation code” controls gene expression during animal development. *Trends in Genetics*, 28(2), pp.78–88.

Alonso, C.R., 2002. Hox Proteins : Sculpting Body Parts by Activating Localised Cell Death Hox proteins shape animal structures by eliciting. *Current Biology*, 12(2), pp.776–778.

Alonso, C.R., 2008. The molecular biology underlying developmental evolution. Cambridge University Press.

Alonso, C.R. & Wilkins, A.S., 2005. The molecular elements that underlie developmental evolution. *Nature reviews. Genetics*, 6(9), pp.709–715.

Alqadah, A., Hsieh, Y.-W. & Chuang, C.-F., 2013. microRNA function in left-right neuronal asymmetry: perspectives from C. elegans. *Frontiers in cellular neuroscience*, 7, pp.1581-1586.

Alvarez-Garcia, I. & Miska, E. a, 2005. MicroRNA functions in animal development and human disease. *Development*, 132(21), pp.4653–4662.

Ambros, V., 2004. The functions of animal microRNAs. *Nature*, 431, pp.350–355.

Ameres, S.L. & Zamore, P.D., 2013. Diversifying microRNA sequence and function. *Nature reviews. Molecular cell biology*, 14(8), pp.475–488.

Aravin, A., Lagos-Quintana, M. & Yalcin, a, 2003. The Small RNA Profile during *Drosophila melanogaster* Development. *Developmental cell*, 5, pp.337–350.

Arif, S. et al., 2013. Evolution of *mir-92a* Underlies Natural Morphological Variation in *Drosophila melanogaster*. *Current Biology*, 23(6), pp.523–528.

Backes, C. et al., 2007. GeneTrail-advanced gene set enrichment analysis. *Nucleic Acids Research*, 35, pp.186–192.

Bandyopadhyay, S. & Mitra, R., 2009. TargetMiner: MicroRNA target prediction with systematic identification of tissue-specific negative examples. *Bioinformatics*, 25(20), pp.2625–2631.

Banerjee-Basu, S. & Baxevanis, a D., 2001. Molecular evolution of the homeodomain family of transcription factors. *Nucleic Acids Res*, 29(15), pp.3258–3269.

Bartel, D.P., 2004. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell*, 116, pp.281–297.

Bartel, D.P., 2009. MicroRNAs: target recognition and regulatory functions. *Cell*, 136(2), pp.215–233.

Bartel, D.P. & Chen, C.-Z., 2004. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nature reviews. Genetics*, 5, pp.396–400.

Bateson, W., 1894. Bateson's Dictionary of Variaton:Materials for the Study of Variation Treated with Especial Regard to Discontinuity in the Origin of Species. *The American Naturalist*, 28, pp.692.

Beachy, P.A., Helfand, S.L. & Hogness, D.S., 1985. Segmental distribution of

bithorax complex proteins during *Drosophila* development. *Nature*, 313, pp.545–551.

Beckervordersandforth, R.M. et al., 2008. Subtypes of glial cells in the *Drosophila* embryonic ventral nerve cord as related to lineage and gene expression. *Mechanisms of Development*, 125, pp.542–557.

Beeman, R.W. et al., 1993. Structure and function of the homeotic gene complex (HOM-C) in the beetle, Tribolium castaneum. *BioEssays*, 15, pp.439–444.

Bejarano, F. et al., 2012. A genome-wide transgenic resource for conditional expression of *Drosophila* microRNAs. *Development*, 139(15), pp.2821–2831.

Bender, W., 2008. MicroRNAs in the *Drosophila* bithorax complex. *Genes & development*, 22(1), pp.14–19.

Bender, W. et al., 1983. Molecular Genetics of the Bithorax Complex in *Drosophila melanogaster*. *Science*, 221, pp.23–29.

Bene, F. Del & Wittbrodt, J., 2005. Cell cycle control by homeobox genes in development and disease. *Seminars in cell & developmental biology*, 16(3), pp.449–460.

Berger, C., Pallavi, S.K., et al., 2005a. A critical role for cyclin E in cell fate determination in the central nervous system of *Drosophila melanogaster*. *Nature cell biology*, 7(1), pp.56–62.

Berger, C., Pallavi, S.K., et al., 2005b. Cyclin E acts under the control of hox-genes as a cell fate determinant in the developing central nervous system. *Cell Cycle*, 4(3), pp.422–425.

Bernstein, E. et al., 2003. Dicer is essential for mouse development. *Nature genetics*, 35(3), pp.215–217.

Betel, D. et al., 2010. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome biology*, 11,

pp.901-914.

Betel, D. et al., 2008. The microRNA.org resource: Targets and expression. *Nucleic Acids Research*, 36, pp.149–153.

Bhat, K.M., 1999. Segment polarity genes in neuroblast formation and identity specification during *Drosophila* neurogenesis. *BioEssays*, 21(6), pp.472–485.

Blair, S.S. et al., 1994. The role of apterous in the control of dorsoventral compartmentalization and PS integrin gene expression in the developing wing of *Drosophila*. *Development*, 120(7), pp.1805–1815.

Boncinelli, E., 1997. Homeobox genes and disease. *Current Opinion in Genetics and Development*, 7(3), pp.331–337.

Bossing, T. et al., 1996. The embryonic central nervous system lineages of *Drosophila melanogaster*. I. Neuroblast lineages derived from the ventral half of the neuroectoderm. *Developmental biology*, 179(1), pp.41–64.

Brand, a H. & Perrimon, N., 1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development*, 118(2), pp.401–415.

Brennecke, J. et al., 2003. bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in *Drosophila*. *Cell*, 113(1), pp.25–36.

Brennecke, J. et al., 2005. Principles of microRNA-target recognition. *PLoS Biology*, 3(3), pp.404–418.

Bridges, C. & Morgan, T., 1923. The third-chromosome group of mutant characters of *Drosophila melanogaster*. *Carniege Institute Publications*, 327, pp.1–251.

Burt, R. & Palka, J., 1982. The central projections of mesothoracic sensory neurons in wild-type *Drosophila* and bithorax mutants. *Developmental biology*, 90(1), pp.99–109.

Carroll, S.B., 1995. Homeotic genes and the evolution of arthropods and chordates. *Nature*. 376, pp.479-485.

Castelli-Gair, J., 1998. Implications of the spatial and temporal regulation of Hox genes on development and evolution. *The International journal of developmental biology*, 42(3), pp.437–444.

Caussinus, E., Colombelli, J. & Affolter, M., 2008. Tip-Cell Migration Controls Stalk-Cell Intercalation during *Drosophila* Tracheal Tube Elongation. *Current Biology*, 18(22), pp.1727–1734.

Chen, C. et al., 2005. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Research*, 33(20), pp.1–9.

Chen, Z. et al., 2012. *miR-92b* regulates Mef2 levels through a negative-feedback circuit during *Drosophila* muscle development. *Development*, 139(19), pp.3543–3552.

Chendrimada, T.P. et al., 2005. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature*, 436, pp.740–744.

Cillo, C. et al., 1999. Homeobox genes and cancer. *Experimental cell research*, 248(1), pp.1–9.

Clark, P.M. et al., 2014. Argonaute CLIP-Seq reveals miRNA targetome diversity across tissue types. *Scientific reports*, 4, pp.59471-59481.

Cohen, S.M., 2010. Macro Roles for MicroRNAs in the Life and Death of Neurons. , pp.91–98.

Cordes, K.R. et al., 2009. miR-145 and miR-143 regulate smooth muscle cell fate and plasticity. *Nature*, 460, pp.705–710.

Coronnello, C. et al., 2012. Novel Modeling of Combinatorial miRNA Targeting Identifies SNP with Potential Role in Bone Density. *PLoS Computational Biology*, 8(12), pp.1–2.

Coronnello, C. & Benos, P. V., 2013. ComiR: Combinatorial microRNA target

prediction tool. *Nucleic acids research*, 41, pp.159–164.

Crickmore, M. a, Ranade, V. & Mann, R.S., 2009. Regulation of Ubx expression by epigenetic enhancer silencing in response to Ubx levels and genetic variation. *PLoS genetics*, 5(9), pp.1-10.

Davis, E. et al., 2005. RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. *Current Biology*, 15, pp.743–749.

Demerec, M., 1950. *Biology of Drosophila*, Wiley. Cold Spring Harbor Laboratory Press.

Doe, C.Q. et al., 1988. Expression and Function of the Segmentation Gene fshi tarazu During *Drosophila* Neurogenesis. *Science*. 239, pp.170-175.

Doe, C.Q., 1992. Molecular markers for identified neuroblasts and ganglion mother cells in the *Drosophila* central nervous system. *Development*, 116(4), pp.855–863.

Duffy, J.B., 2002. GAL4 system in *Drosophila*: a fly geneticist's Swiss army knife. *Genesis*, 34, pp.1–15.

Dweep, H. et al., 2011. MiRWalk - Database: Prediction of possible miRNA binding sites by " walking" the genes of three genomes. *Journal of Biomedical Informatics*, 44(5), pp.839–847.

Elcheva, I. et al., 2009. CRD-BP Protects the Coding Region of TrCP1 mRNA from miR-183-Mediated Degradation. *Molecular Cell*, 35(2), pp.240–246.

Enright, A.J. et al., 2003. MicroRNA targets in *Drosophila*. *Genome biology*, 5(1), pp.1-14.

Enright, A.J., prediction of microRNA targets. *Methods Mol Biol*, 667, pp.283–295.

Eulalio, A., Huntzinger, E. & Izaurralde, E., 2008. Getting to the root of miRNA-mediated gene silencing. *Cell*, 132(1), pp.9–14.

Fernandes, J. et al., 1994. Muscle development in the four-winged *Drosophila*

and the role of the Ultrabithorax gene. *Current biology : CB*, 4(11), pp.957–964.

Filipowicz, W., Bhattacharyya, S.N. & Sonenberg, N., 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nature reviews. Genetics*, 9(2), pp.102–114.

Fontana, L. et al., 2007. MicroRNAs 17-5p-20a-106a control monocytopoiesis through AML1 targeting and M-CSF receptor upregulation. *Nature cell biology*, 9(7), pp.775–787.

Forman, J.J. & Coller, H. a., 2010. The code within the code: MicroRNAs target coding regions. *Cell Cycle*, 9, pp.1533–1541.

Freeman, M.R. & Doherty, J., 2006. Glial cell biology in *Drosophila* and vertebrates. *Trends in Neurosciences*, 29(2), pp.82–90.

Gaidatzis, D. et al., 2007. Inference of miRNA targets using evolutionary conservation and pathway analysis. *BMC bioinformatics*, 8, pp.1-22.

Garaulet, D.L. et al., 2014. Homeotic Function of *Drosophila* Bithorax-Complex miRNAs Mediates Fertility by Restricting Multiple Hox Genes and TALE Cofactors in the CNS. *Developmental Cell*, 29(6), pp.635–648.

Gaunt, S.J., 2015. The significance of Hox gene collinearity. *The International Journal of Developmental Biology*, 59, pp159-170.

Gennarino, V.A. et al., 2011. HOCTAR database: A unique resource for microRNA target prediction. *Gene*, 480, pp.51–58.

Giammartino, D.C. Di, Nishida, K. & Manley, J.L., 2011. Mechanisms and consequences of alternative polyadenylation. *Molecular cell*, 43(6), pp.853–866.

Griffiths-Jones, S. et al., 2011. MicroRNA evolution by arm switching. *EMBO reports*, 12(2), pp.172–177.

Grimson, A. et al., 2007. MicroRNA Targeting Specificity in Mammals:

Determinants beyond Seed Pairing. *Molecular Cell*, 27, pp.91–105.

Grishok, A. et al., 2001. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell*, 106(1), pp.23–34.

Grün, D. et al., 2005. microRNA target predictions across seven *Drosophila* species and comparison to mammalian targets. *PLoS computational biology*, 1(1), pp.51-66.

Gummalla, M. et al., 2014. Hox gene regulation in the central nervous system of *Drosophila*. *Frontiers in cellular neuroscience*, 8(96), pp.1-12.

Hafner, M. et al., 2010. Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP. *Cell*, 141(1), pp.129–141.

Hammell, M. et al., 2008. mirWIP: microRNA target prediction based on microRNA-containing ribonucleoprotein-enriched transcripts. *Nature methods*, 5, pp.813–819.

Han, J. et al., 2004. The Drosha-DGCR8 complex in primary microRNA processing. *Genes and Development*, 18(24), pp.3016–3027.

Hartenstein, V., 1993. Atlas of *Drosophila* Development, Cold Spring Harbor Laboratory Press.

He, X., Yan, Y. & Eberhart, johann k., 2012. miR-196 regulates axial patterning and pectoral appendage initiation. *Developmental Biology*, 29(6), pp.997–1003.

Heitzler, P. et al., 1996. Genes of the Enhancer of split and achaete-scute complexes are required for a regulatory loop between Notch and Delta during lateral signalling in *Drosophila*. *Development*, 122(1), pp.161–171.

Heitzler, P. & Simpson, P., 1991. The choice of cell fate in the epidermis of *Drosophila*. *Cell*, 64(6), pp.1083–1092.

Held, L.I.J., 2002. Imaginal Discs: The Genetic and Cellular Logic of Pattern

Formation (CHAPTER EIGHT Homeosis), Cambridge University Press.

Hersh, B.M. et al., 2007. The UBX-regulated network in the haltere imaginal disc of D. melanogaster. *Developmental Biology*, 302(2), pp.717–727.

Hirth, F., Hartmann, B. & Reichert, H., 1998. Homeotic gene action in embryonic brain development of *Drosophila*. *Development*, 125(9), pp.1579–1589.

Holland, L.Z. et al., 2013. Evolution of bilaterian central nervous systems: a single origin? *EvoDevo*, 4(1), pp.1-20.

Hsu, S. Da et al., 2008. miRNAMap 2.0: Genomic maps of microRNAs in metazoan genomes. *Nucleic Acids Research*, 36, pp.165–169.

Huang, J.C., Morris, Q.D. & Frey, B.J., 2007. Bayesian inference of MicroRNA targets from sequence and expression data. *Journal of computational biology : a journal of computational molecular cell biology*, 14(5), pp.550–563.

Hughes, C.L. & Kaufman, T.C., 2002. Hox genes and the evolution of the arthropod body plan. *Evolution & development*, 4(6), pp.459–499.

Hui, J.H.L. et al., 2013. Structure, evolution and function of the bi-directionally transcribed iab-4/iab-8 microRNA locus in arthropods. *Nucleic Acids Research*, 41(5), pp.3352–3361.

Hwang, W.-L. et al., 2014. MicroRNA-146a directs the symmetric division of Snail-dominant colorectal cancer stem cells. *Nature cell biology*, 16(3), pp.268–280.

Ingham, P.W. & Martinez-Arias, a, 1986. The correct activation of Antennapedia and bithorax complex genes requires the fushi tarazu gene. *Nature*, 324(6097), pp.592–597.

Innis, J.W. et al., 2002. A HOXA13 allele with a missense mutation in the homeobox and a dinucleotide deletion in the promoter underlies Guttmacher syndrome. *Human mutation*, 19(5), pp.573–574.

Irvine, K.D., Helfand, S.L. & Hogness, D.S., 1991. The large upstream control region of the *Drosophila* homeotic gene Ultrabithorax. *Development*, 111(2), pp.407–424.

Ito, K., Urban, J. & Technau, G.M., 1995. Distribution, classification, and development of *Drosophila* glial cells in the late embryonic and early larval ventral nerve cord. *Roux's Archives of Developmental Biology*, 204(5), pp.284–307.

Ivey, K.N. & Srivastava, D., 2010. MicroRNAs as regulators of differentiation and cell fate decisions. *Cell Stem Cell*, 7(1), pp.36–41.

Jarvis, E., Bruce, H.S. & Patel, N.H., 2012. Evolving specialization of the arthropod nervous system. *Proceedings of the National Academy of Sciences of the United States of America*, 109, pp.10634–10639.

Jenett, A. et al., 2012. A GAL4-Driver Line Resource for *Drosophila* Neurobiology. *Cell Reports*, 2(4), pp.991–1001.

Jiménez, F. & Campos-Ortega, J. a, 1990. Defective neuroblast commitment in mutants of the achaete-scute complex and adjacent genes of D. melanogaster. *Neuron*, 5(1), pp.81–89.

John, B. et al., 2004. Human microRNA targets. *PLoS Biology*, 2(11).pp.e363.

Kaschula, R., 2013. The regulation of Hox genes by microRNAs during *Drosophila* development. *Phd thesis*.

Kaufman, T.C., Lewis, R. & Wakimoto, B., 1980. Cytogenetic analysis of chromosome 3 in *Drosophila melanogaster*: The homoeotic gene complex in polytene chromosome interval 84A-B. *Genetics*, 94(1), pp.115–133.

Kennison, J. a, 1993. Transcriptional activation of *Drosophila* homeotic genes from distant regulatory elements. *Trends in genetics : TIG*, 9(3), pp.75–79.

Kertesz, M. et al., 2007. The role of site accessibility in microRNA target recognition. *Nature genetics*, 39(10), pp.1278–1284.

Kheradpour, P. et al., 2007. Reliable prediction of regulator targets using 12 *Drosophila* genomes. *Genome Research*, 17, pp.1919–1931.

Khila, A., Abouheif, E. & Rowe, L., 2009. Evolution of a novel appendage ground plan in water striders is driven by changes in the Hox gene Ultrabithorax. *PLoS Genetics*, 5(7), pp.1–9.

Kim, V.N., 2004. MicroRNA precursors in motion: Exportin-5 mediates their nuclear export. *Trends in Cell Biology*, 14(4), pp.156–159.

Kim, V.N., Han, J. & Siomi, M.C., 2009. Biogenesis of small RNAs in animals. *Nature reviews. Molecular cell biology*, 10(2), pp.126–139.

King, M.-C., Wilson, A.C. & Url, S., 1975. Evolution at Two Levels in Humans and Chimpanzees. *Science*, 188, pp.107–116.

Kiriakidou, M. et al., 2004. A combined computational-experimental approach predicts human microRNA targets. *Genes and Development*, 18, pp.1165–1178.

Kloosterman, W.P. & Plasterk, R.H. a, 2006. The diverse functions of microRNAs in animal development and disease. *Developmental cell*, 11(4), pp.441–450.

Kornberg, T.B. & Tabata, T., 1993. Segmentation of the *Drosophila* embryo. *Current opinion in genetics & development*, 3(4), pp.585–594.

Kornfeld, K. et al., 1989. Structure and expression of a family of Ultrabithorax mRNAs generated by alternative splicing and polyadenylation in *Drosophila*. *Genes & development*, 3(2), pp.243–258.

Kosik, K.S., 2009. MicroRNAs tell an evo-devo story. *Nature reviews. Neuroscience*, 10(10), pp.754–759.

Kourakis, M.J. et al., 1997. Conserved anterior boundaries of Hox gene expression in the central nervous system of the leech Helobdella. *Developmental biology*, 190(2), pp.284–300.

Kozomara, A. & Griffiths-Jones, S., 2014. MiRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research*, 42(1), pp.68–73.

Krek, A. et al., 2005. Combinatorial microRNA target predictions. *Nature genetics*, 37(5), pp.495–500.

Krol, J., Loedige, I. & Filipowicz, W., 2010. The widespread regulation of microRNA biogenesis, function and decay. *Nature reviews. Genetics*, 11(9), pp.597–610.

Krüger, J. & Rehmsmeier, M., 2006. RNAhybrid: MicroRNA target prediction easy, fast and flexible. *Nucleic Acids Research*, 34(16), pp.451–454.

Lagos-Quintana, M. et al., 2001. Identification of novel genes coding for small expressed RNAs. *Science*, 294, pp.853–858.

Lai, E.C. et al., 2005. Pervasive regulation of *Drosophila* Notch target genes by GY-box-, Brd-box-, and K-box-class microRNAs. *Genes & Development*, pp.1067–1080.

Lau, N.C. et al., 2001. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. *Science*, 294, pp.858–862.

Lecuit, T. et al., 1996. Two distinct mechanisms for long-range patterning by Decapentaplegic in the *Drosophila* wing. *Nature*, 381, pp.387–393.

Lee, R.C., 1993. The C . elegans Heterochronic Gene lin-4 Encodes Small RNAs with Antisense Complementarity to II-14. *Cell*, 75, pp.843–854.

Lee, R.C. & Ambros, V., 2001. An extensive class of small RNAs in Caenorhabditis elegans. *Science*, 294, pp.862–864.

Lee, Y. et al., 2004. MicroRNA genes are transcribed by RNA polymerase II. *Embo Journal*, 23(20), pp.4051–4060.

Lee, Y.S. et al., 2004. Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell*, 117(1), pp.69–81.

Lewis, B.P. et al., 2003. Prediction of Mammalian MicroRNA Targets. *Cell*, 115, pp.787–798.

Lewis, B.P., Burge, C.B. & Bartel, D.P., 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, 120(1), pp.15–20.

Lewis, E.B., 1978. A gene complex controlling segmentation in *Drosophila*. *Nature*, 276(5688), pp.565–570.

Lewis, E.B., 1951. *Pseudoallelism and Gene Evolution*, Cold Spring Harbor Laboratory Press.

Lewis, R. a. et al., 1980a. Genetic analysis of the antennapedia gene complex (ant-C) and adjacent chromosomal regions of *Drosophila melanogaster*. II. Polytene chromosome segments 84A-84B1,2. *Genetics*, 95(2), pp.383–397.

Lewis, R.A. et al., 1980b. Genetic Analysis of the Antennapedia Gene Complex (Ant-C) and Adjacent Chromosomal Regions of *Drosophila melanogaster*. I. Polytene Chromosome Segments 84b-D. *Genetics*, 95, pp.367–381.

Li, H.H. et al., 2014. A GAL4 Driver Resource for Developmental and Behavioral Studies on the Larval CNS of *Drosophila*. *Cell Reports*, 8(3), pp.897–908.

Lin, C.-C. et al., 2011. A KLF4-miRNA-206 autoregulatory feedback loop can promote or inhibit protein translation depending upon cell context. *Molecular and cellular biology*, 31(12), pp.2513–2527.

Little, J.W. et al., 1990. Effect of abx, bx and pbx mutations on expression of homeotic genes in *Drosophila* larvae. *Genetics*, 124(4), pp.899–908.

Liu, J. et al., 2004. Argonaute2 is the catalytic engine of mammalian RNAi. *Science*, 305, pp.1437–1441.

Liu, N. et al., 2012. The microRNA miR-34 modulates ageing and neurodegeneration in *Drosophila*. *Nature*, 482, pp.519–523.

Maeda, R.K. & Karch, F., 2006. The ABC of the BX-C: the bithorax complex explained. *Development*, 133(8), pp.1413–1422.

Majoros, W.H. & Ohler, U., 2007. Spatial preferences of microRNA targets in 3' untranslated regions. *BMC genomics*, 8, pp.1521-1529.

Mallo, M. & Alonso, C.R., 2013. The regulation of Hox gene expression during animal development. *Development*, 140(19), pp.3951–3963.

Mann, R.S., 1994. Engrailed-mediated repression of Ultrabithorax is necessary for the parasegment 6 identity in *Drosophila*. *Development*, 120(11), pp.3205–3212.

Manning, L. et al., 2012. A resource for manipulating gene expression and analyzing cis-regulatory modules in the *Drosophila* CNS. *Cell reports*, 2(4), pp.1002–1013.

Marin, E.C. et al., 2012. Ultrabithorax confers spatial identity in a context-specific manner in the *Drosophila* postembryonic ventral nervous system. *Neural Development*, 7(1), pp.1-15.

McGinnis, W. et al., 1984a. A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. *Cell*, 37(2), pp.403–408.

McGinnis, W. et al., 1984b. Molecular cloning and chromosome mapping of a mouse DNA sequence homologous to homeotic genes of *Drosophila*. *Cell*, 38(3), pp.675–680.

McGlinn, E. et al., 2009. In ovo application of antagomiRs indicates a role for miR-196 in patterning the chick axial skeleton through Hox gene regulation. *Proceedings of the National Academy of Sciences of the United States of America*, 106(44), pp.18610–18615.

Meyer, A., 1998. hox gene variation and evolution. *Nature*, 17(4), pp.471–480.

Meza-Sosa, K.F. et al., 2012. Role of microRNAs in central nervous system development and pathology. *Journal of neuroscience research*, 90(1),

pp.1–12.

Miguel-Aliaga, I. & Thor, S., 2004. Segment-specific prevention of pioneer neuron apoptosis by cell-autonomous, postmitotic Hox gene activity. *Development*, 131(24), pp.6093–6105.

Miranda, K.C. et al., 2006. A Pattern-Based Method for the Identification of MicroRNA Binding Sites and Their Corresponding Heteroduplexes. *Cell*, 126, pp.1203–1217.

Moore, M.J., 2005. From birth to death: the complex lives of eukaryotic mRNAs. *Science*, 309, pp.1514–1518.

Morata, G., 2001. How *Drosophila* appendages develop. *Nature reviews. Molecular cell biology*, 2(2), pp.89–97.

Mortazavi, A. et al., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods*, 5(7), pp.621–628.

Muragaki, Y. et al., 1996. Altered growth and branching patterns in synpolydactyly caused by mutations in HOXD13. *Science*, 272, pp.548–551.

Nairz, K. et al., 2006. Overgrowth caused by misexpression of a microRNA with dispensable wild-type function. *Developmental biology*, 291(2), pp.314–324.

Nam, S. et al., 2008. miRGator: An integrated system for functional annotation of microRNAs. *Nucleic Acids Research*, 36, pp.159–164.

Navas, L. de et al., 2006. A simple and efficient method to identify replacements of P-lacZ by P-Gal4 lines allows obtaining Gal4 insertions in the bithorax complex of *Drosophila*. *Mechanisms of development*, 123(11), pp.860–867.

Navas, L.F. de et al., 2011. Integration of RNA processing and expression level control modulates the function of the *Drosophila* Hox gene Ultrabithorax during adult development. *Development*, 138(1), pp.107–116.

Navas, L.F. de, Garaulet, D.L. & Sánchez-Herrero, E., 2006. The ultrabithorax

Hox gene of *Drosophila* controls haltere size by regulating the Dpp pathway. *Development*, 133(22), pp.4495–4506.

Nielsen, C.B. et al., 2007. Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *RNA*, 13(11), pp.1894–1910.

Nilsen, T.W., 2007. Mechanisms of microRNA-mediated gene regulation in animal cells. *Trends in genetics : TIG*, 23(5), pp.243–249.

O'Connor, M.B. et al., 1988. Alternative RNA products from the Ultrabithorax domain of the bithorax complex. *The EMBO journal*, 7(2), pp.435–445.

Ohlen, T. von & Doe, C.Q., 2000. Convergence of dorsal, dpp, and egfr signaling pathways subdivides the *Drosophila* neuroectoderm into three dorsal-ventral columns. *Developmental biology*, 224(2), pp.362–372.

Okamura, K. et al., 2004. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes and Development*, 18(14), pp.1655–1666.

Okamura, K. et al., 2008. The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution. *Nature structural & molecular biology*, 15(4), pp.354–363.

Okamura, K., Liu, N. & Lai, E.C., 2009. Distinct Mechanisms for MicroRNA Strand Selection by *Drosophila* Argonautes. *Molecular Cell*, 36(3), pp.431–444.

Orphanides, G. & Reinberg, D., 2002. A unified theory of gene expression. *Cell*, 108, pp.439–451.

Ørom, U.A., Nielsen, F.C. & Lund, A.H., 2008. MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Molecular cell*, 30(4), pp.460–471.

Paraskevopoulou, M.D. et al., 2013. DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. *Nucleic acids research*, 41, pp.169–173.

Paro, R., 1990. Imprinting a determined state into the chromatin of *Drosophila*. *Trends in Genetics*, 6(12), pp.416–421.

Pasquinelli, a E. et al., 2000. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature*, 408, pp.86–89.

Patraquim, P., Warnefors, M. & Alonso, C.R., 2011. Evolution of hox post-transcriptional regulation by alternative polyadenylation and MicroRNA modulation within 12 *Drosophila* genomes. *Molecular Biology and Evolution*, 28(9), pp.2453–2460.

Pavlopoulos, A. & Akam, M., 2011. Hox gene Ultrabithorax regulates distinct sets of target genes at successive stages of *Drosophila* haltere morphogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 108(7), pp.2855–2860.

Pearson, J.C., Lemons, D. & McGinnis, W., 2005. Modulating Hox gene functions during animal body patterning. *Nature reviews. Genetics*, 6(12), pp.893–904.

Picao-Osorio, J. et al., 2015. MicroRNA-encoded behavior in *Drosophila*. *Science*, 350, pp.1–17.

Pietro, M. di et al., 2012. Evidence for a functional role of epigenetically regulated midcluster HOXB genes in the development of Barrett esophagus. *Proceedings of the National Academy of Sciences*, 109(23), pp.9077–9082.

Pirrotta, V., 1997. Chromatin-silencing mechanisms in *Drosophila* maintain patterns of gene expression. *Trends in Genetics*, 13(8), pp.314–318.

Procino, A. & Cillo, C., 2013. The HOX genes network in metabolic diseases. *Cell Biology International*, 37(11), pp.1145–1148.

Prokop, a & Technau, G.M., 1994. Early tagma-specific commitment of *Drosophila* CNS progenitor NB1-1. *Development*, 120(9), pp.2567–2578.

Qian, S., Capovilla, M. & Pirrotta, V., 1993. Molecular mechanisms of pattern formation by the BRE enhancer of the Ubx gene. *The EMBO journal*,

12(10), pp.3865–3877.

Quinonez, S.C. & Innis, J.W., 2014. Human HOX gene disorders. *Molecular Genetics and Metabolism*, 111(1), pp.4–15.

Rajewsky, N., 2006. microRNA target predictions in animals. *Nature genetics*, 38, pp.8–13.

Raman, V. et al., 2000. Compromised HOXA5 function can limit p53 expression in human breast tumours. *Nature*, 405, pp.974–978.

Reczko, M. et al., 2012. Functional microRNA targets in protein coding sequences. *Bioinformatics*, 28(6), pp.771–776.

Reed, H.C. et al., 2010. Alternative splicing modulates Ubx protein function in *Drosophila melanogaster*. *Genetics*, 184(3), pp.745–58.

Rehmsmeier, M. et al., 2004. Fast and effective prediction of microRNA / target duplexes. *RNA*, 10, pp.1507–1517.

Reinhart, B.J. et al., 2000. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature*, 403, pp.901–906.

Reinitz, J. & Levine, M., 1990. Control of the initiation of homeotic gene expression by the gap genes giant and tailless in *Drosophila*. *Developmental Biology*, 140(1), pp.57–72.

Rennie, W. et al., 2014. STarMir: A web server for prediction of microRNA binding sites. *Nucleic Acids Research*, 42, pp.114–118.

Roch, F. & Akam, M., 2000. Ultrabithorax and the control of cell morphology in *Drosophila* halteres. *Development*, 127(1), pp.97–107.

Roch, F., Alonso, C.R. & Akam, M., 2003. *Drosophila* miniature and dusky encode ZP proteins required for cytoskeletal reorganisation during wing morphogenesis. *Journal of cell science*, 116, pp.1199–1207.

Rogulja-Ortmann, A. et al., 2014. The RNA-binding protein ELAV regulates Hox RNA processing, expression and function within the *Drosophila* nervous

system. *Development*, 141(10), pp.2046–2056.

Rogulja-Ortmann, A., Renner, S. & Technau, G.M., 2008. Antagonistic roles for Ultrabithorax and Antennapedia in regulating segment-specific apoptosis of differentiated motoneurons in the *Drosophila* embryonic central nervous system. *Development*, 135(20), pp.3435–3445.

Ronshaugen, M. et al., 2005. The *Drosophila* microRNA iab-4 causes a dominant homeotic transformation of halteres to wings. *Genes & development*, 19(24), pp.2947–2952.

Rozowski, M. & Akam, M., 2002. Hox gene control of segment-specific bristle patterns in *Drosophila*. *Genes & development*, 16(9), pp.1150–1162.

Ruby, J.G. et al., 2007. Evolution , biogenesis , expression , and target predictions of a substantially expanded set of *Drosophila* microRNAs. *Genome research*, 17, pp.1850–1864.

Rusinov, V. et al., 2005. MicroInspector: A web tool for detection of miRNA binding sites in an RNA sequence. *Nucleic Acids Research*, 33, pp.696–700.

Sætrom, P. et al., 2007. Distance constraints between microRNA target sites dictate efficacy and cooperativity. *Nucleic Acids Research*, 35(7), pp.2333–2342.

Saito, T. & Sætrom, P., 2010. MicroRNAs - targeting and target prediction. *New Biotechnology*, 27(3), pp.243–249.

Sales, G. et al., 2010. Magia, a web-based tool for miRNA and genes integrated analysis. *Nucleic Acids Research*, 38, pp.1–8.

Sandberg, R. et al., 2008. Proliferating cells express mRNAs with shortened 3’ UTRs and fewer microRNA target sites. *Science*, 320, pp.1643–1647.

Saugstad, J.A., 2010. MicroRNAs as effectors of brain function with roles in ischemia and injury , neuroprotection , and neurodegeneration. *Journal of Cerebral Blood Flow & Metabolism*, 30(9), pp.1564–1576.

Schertel, C. et al., 2012. Functional characterization of *Drosophila* microRNAs by a novel in vivo library. *Genetics*, 192(4), pp.1543–1552.

Schirle, N.T., Sheu-gruttadauria, J. & Macrae, I.J., 2014. Structural basis for microRNA targeting. *Science*, 346(6209), pp.608–613.

Schmid, A, Chiba, A & Doe, C.Q., 1999. Clonal analysis of *Drosophila* embryonic neuroblasts: neural cell types, axon projections and muscle targets. *Development*, 126(21), pp.4653–4689.

Schmidt, H. et al., 1997. The embryonic central nervous system lineages of *Drosophila melanogaster*. II. Neuroblast lineages derived from the dorsal part of the neuroectoderm. *Developmental biology*, 189(2), pp.186–204.

Schuldiner, O. et al., 2008. piggyBac-based mosaic screen identifies a postmitotic function for cohesin in regulating developmental axon pruning. *Developmental cell*, 14(2), pp.227–238.

Sethupathy, P., Megraw, M. & Hatzigeorgiou, A.G., 2006. A guide through present computational approaches for the identification of mammalian microRNA targets. *Nature methods*, 3(11), pp.881–886.

Shenoy, A. & Blelloch, R.H., 2014. Regulation of microRNA function in somatic stem cell proliferation and differentiation. *Nature Reviews Molecular Cell Biology*, 15(9), pp.565–576.

Simon, J., 1995. Locking in stable states of gene expression: transcriptional control during *Drosophila* development. *Curr Opin Cell Biol*, 7(3), pp.376–385.

Skeath, J.B., 1998. The *Drosophila* EGF receptor controls the formation and specification of neuroblasts along the dorsal-ventral axis of the *Drosophila* embryo. *Development*, 125(17), pp.3301–3312.

Skeath, J.B. & Carroll, S.B., 1992. Regulation of proneural gene expression and cell fate during neuroblast segregation in the *Drosophila* embryo. *Development*, 114(4), pp.939–946.

Skeath, J.B. & Thor, S., 2003. Genetic control of *Drosophila* nerve cord development. *Current Opinion in Neurobiology*, 13(1), pp.8–15.

Slack, F.J. et al., 2000. The lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. *Molecular cell*, 5(4), pp.659–669.

Smolik-Utlaut, S.M., 1990. Dosage requirements of Ultrabithorax and bithoraxoid in the determination of segment identity in *Drosophila melanogaster*. *Genetics*, 124(2), pp.357–366.

Stark, A. et al., 2008. A single Hox locus in *Drosophila* produces functional microRNAs from opposite DNA strands. *Genes & development*, 22(1), pp.8–13.

Stark, A. et al., 2005. Animal microRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell*, 123, pp.1133–1146.

Stark, A. et al., 2003. Identification of *Drosophila* MicroRNA targets. *PLoS biology*, 1(3), p.397-409.

Steffen, P. a & Ringrose, L., 2014. What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. *Nature reviews. Molecular cell biology*, 15(5), pp.340–356.

Stern, D.L., Road, T.C. & Cb, C., 1998. A role of ultrabithorax in morphological differences between *Drosophila* species. *Nature*, 396, pp.463–467.

Stork, T. et al., 2008. Organization and function of the blood-brain barrier in *Drosophila*. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 28(3), pp.587–597.

Sullivan, W., Ashburner, M. & Hawley, R.S., 2000. *Drosophila protocol*, Cold Spring Harbor Laboratory Press.

Sun, M. et al., 2013. HMGA2/TET1/HOXA9 signaling pathway regulates breast cancer growth and metastasis. *Proceedings of the National Academy of Sciences of the United States of America*, 110(24), pp.9920–9925.

Suska, A., Miguel-Aliaga, I. & Thor, S., 2011. Segment-specific generation of *Drosophila* Capability neuropeptide neurons by multi-faceted Hox cues. *Developmental Biology*, 353(1), pp.72–80.

Szuplewski, S. et al., 2012. MicroRNA transgene over expression complements deficiency-based modifier screens in *Drosophila*. *Genetics*, 190(2), pp.617–626.

Tan, C.L. et al., 2013. MicroRNA-128 governs neuronal excitability and motor behavior in mice. *Science*, 342, pp.1254–1258.

Tay, Y. et al., 2008. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature*, 455, pp.1124–1128.

Technau, G.M., Berger, C. & Urbach, R., 2006. Generation of cell diversity and segmental pattern in the embryonic central nervous system of *Drosophila*. *Developmental dynamics : an official publication of the American Association of Anatomists*, 235(4), pp.861–869.

Thadani, R. & Tammi, M.T., 2006. MicroTar: predicting microRNA targets from RNA duplexes. *BMC bioinformatics*, 7, pp.1-0.

Thompson, B.J. & Cohen, S.M., 2006. The Hippo Pathway Regulates the bantam microRNA to Control Cell Proliferation and Apoptosis in *Drosophila*. *Cell*, 126, pp.767–774.

Thomsen, S. et al., 2010. Developmental RNA processing of 3'UTRs in Hox mRNAs as a context-dependent mechanism modulating visibility to microRNAs. *Development*. 137(17), pp.2951-2960.

Tyler, D.M. et al., 2008. Functionally distinct regulatory RNAs generated by bidirectional transcription and processing of microRNA loci. *Genes & development*, 22(1), pp.26–36.

Udolph, G. et al., 1993. A common precursor for glia and neurons in the embryonic CNS of *Drosophila* gives rise to segment-specific lineage variants. *Development*, 118(3), pp.765–775.

Vasudevan, S., 2012. Posttranscriptional upregulation by microRNAs. *Wiley interdisciplinary reviews. RNA*, 3(3), pp.311–330.

Wahid, F. et al., 2010. MicroRNAs: Synthesis, mechanism, function, and recent clinical trials. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1803(11), pp.1231–1243.

Wang, W. X. et al., 2010. Individual microRNAs (miRNAs) display distinct mRNA targeting “rules”. *RNA biology*, 7, pp.373–380.

Warren, R.W. et al., 1994. Evolution of homeotic gene regulation and function in flies and butterflies. *Nature*, 372, pp.458–461.

Weatherbee, S.D. et al., 1998. Ultrabithorax regulates genes at several levels of the wing-patterning hierarchy to shape the development of the *Drosophila* haltere. *Genes and Development*, 12(10), pp.1474–1482.

Wellik, D.M., 2007. Hox patterning of the vertebrate axial skeleton. *Developmental Dynamics*, 236(9), pp.2454–2463.

White, R. a & Lehmann, R., 1986. A gap gene, hunchback, regulates the spatial expression of Ultrabithorax. *Cell*, 47(2), pp.311–321.

White, R. a & Wilcox, M., 1985. Distribution of Ultrabithorax proteins in *Drosophila*. *The EMBO journal*, 4(8), pp.2035–2043.

White, R. a & Wilcox, M., 1984. Protein products of the bithorax complex in *Drosophila*. *Cell*, 39(1), pp.163–171.

Wienholds, E. et al., 2003. The microRNA-producing enzyme Dicer1 is essential for zebrafish development. *Nature genetics*, 35(3), pp.217–218.

Wightman, B., Ha, I. & Ruvkun, G., 1993. Posttranscriptional regulation of the heterochronic gene lin- 14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell*, 75, pp.855–862.

Wilhelm, B.T. et al., 2008. Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature*, 453, pp.1239–1243.

Witkos, T.M., Koscianska, E. & Krzyzosiak, W.J., 2011. Practical Aspects of microRNA Target Prediction. *Current molecular medicine*, 11(2), pp.93–109.

Woltering, J.M. & Durston, A.J., 2008. MiR-10 Represses HoxB1a and HoxB3a in Zebrafish. *PLoS ONE*, 3(1), pp.1-13.

Wong, N. & Wang, X., 2014. miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Research*, 43(1), pp.146–152.

Wu, Y.C. et al., 2012. Let-7-Complex MicroRNAs Regulate the Temporal Identity of *Drosophila* Mushroom Body Neurons via chinmo. *Developmental Cell*, 23(1), pp.202–209.

Xiao, F. et al., 2009. miRecords: An integrated resource for microRNA-target interactions. *Nucleic Acids Research*, 37, pp.105–110.

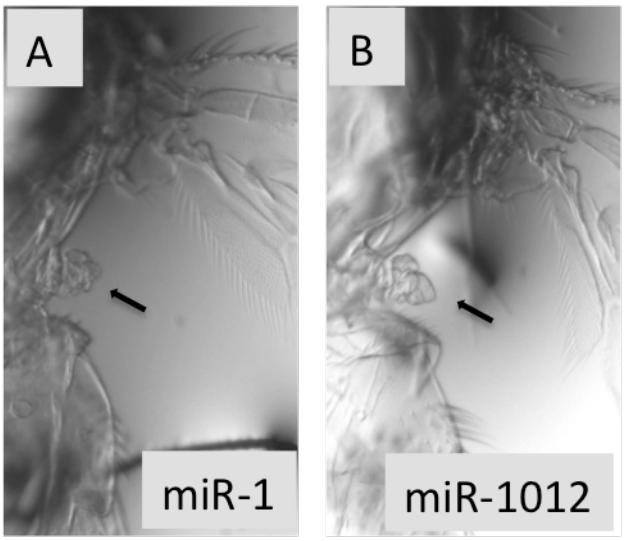
Xu, J. & Wong, C., 2008. A computational screen for mouse signaling pathways targeted by microRNA clusters. *RNA*, 14(7), pp.1276–1283.

Yekta, S., Shih, I.-H. & Bartel, D.P., 2004. MicroRNA-directed cleavage of HOXB8 mRNA. *Science*, 304, pp.594–596.

Zeng, Y. & Cullen, B.R., 2005. Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. *Journal of Biological Chemistry*, 280(30), pp.27595–27603.

Zhou, H. & Rigoutsos, I., 2014. MiR-103a-3p targets the 5 ′ UTR of GPRC5A in pancreatic cells. *RNA*, 20, pp.1–9.

APPENDIX



Appendix-Figure 1 Loss of the capitellum of the haltere caused by misexpression of different Class III miRNAs.

All halteres shown here are female halteres. Ectopic expression of *miR-1* and *miR-1012* by *Ubx*-Gal4^{LDN} cause the loss of the capitellum of the haltere. Black arrowheads point to the position of the haltere.

Appendix-Table 1 miRNAs used for TaqMan RT-PCR quantification.

microRNAs	Haltere (FPKM)	Expression	miRNA Classes
mir-252-5p	6.67E+11	High	I
mir-304-5p	66412300000	Medium	I
mir-92a-3p	1.25E+14	Very High	I
mir-iab-4-5p	9387530000	Low	I
mir-282-5p	7.00E+12	High	II
mir-980-3p	37550100000	Low	II
mir-1012-3p	5.19E+12	High	III
mir-124-3p	33206200000	Low	III
mir-1006-3p	39847400000	Medium	IV
mir-137-3p	65712700000	Medium	IV
mir-190-5p	2.12E+11	Medium	IV
mir-9a-5p	8.53E+14	Very High	IV
mir-313-3p	9387530000	Low	lethal
mir-375-3p	2.09E+12	High	lethal
let-7-5p	1.90E+13	Very High	lethal

According to RNA-seq data (collected by our previous lab member Dr. Richard Kaschula), all the miRNAs were listed by overall abundance with the haltere tissue. Then all sequenced miRNAs were divided into quartile groups - 0-25%, 26-50%, 51-75% and 76-100% based on the haltere abundance values. 76-100%: 'Very High', 51-75%: 'High', 26-50%: 'Medium', 0-25%: 'Low'. The miRNA Classes is classified according to the severity of the haltere morphological changes by ectopic expression of miRNAs as described in Figure 3.3-3.6.

Appendix-Table 2 The penetrance of the genetic screen.

microRNA	microRNA class	Penetrance
miR-10	1	100%
miR-252	1	98%
miR-303	1	100%
miR-304	1	100%
miR-310	1	100%
miR-311	1	100%
miR-312	1	100%
miR-92a	1	100%
miR-92b	1	100%
miR-iab-4	1	100%
miR-12	2	61.7%
miR-133	2	100%
miR-278	2	81.3%
miR-282	2	100%
miR-285	2	92%
miR-31b	2	100%
miR-33	2	79.6%
miR-6-1, 6-2, 6-3	2	66.7%
miR-7	2	86.3%
miR-79	2	63.3%
miR-8	2	100%
miR-927	2	84.2%
miR-958	2	75.7%
miR-980	2	100.00%
miR-982	2	100%
miR-995	2	65.5%
miR-bft (miR-263a)	2	100%
bantam	2	100%
miR-274	3	100%
miR-1	3	100%
miR-1000	3	100%
miR-1012	3	100%
miR-124	3	100%
miR-14	3	100%
miR-184	3	100%
miR-210	3	100%

miR-263b	3	100%
miR-276a	3	100%
miR-276b	3	100%
miR-279	3	100%
miR-284	3	100%
miR-306, 9b, 9c, 79	3	100.00%
miR-307a	3	100%
miR-318	3	98.4%
miR-34	3	100%
miR-4966,975,976,977	3	97.9%
miR-5,4,286	3	100%
miR-932	3	100%
miR-963,964	3	72.3%
miR-964	3	100%
miR-976	3	98%
miR-984	3	100%
miR-985	3	100%
miR-989	3	100%
miR-999	3	100%
miR-9b	3	100%
miR-9c	3	100%
miR-100	4	100%
miR-1001	4	100%
miR-1003	4	98.3%
miR-1004	4	100%
miR-1006	4	88.8%
miR-1007	4	100%
miR-1009	4	100%
miR-1010	4	100%
miR-1011	4	100%
miR-1013	4	100%
miR-1015	4	100%
miR-1017	4	100%
miR-137	4	98.3%
miR-190	4	100%
miR-275	4	100%
miR-280	4	100%
miR-281	4	100%
miR-281-1, 281-2	4	100%

miR-281-1	4	100%
miR-286	4	100%
miR-2b-1	4	100%
miR-305	4	100%
miR-305, 275	4	100%
miR-308	4	97.3%
miR-308	4	100%
miR-309	4	100%
miR-31a	4	100%
miR-929	4	100%
miR-954	4	100%
miR-955	4	100%
miR-956	4	100%
miR-961	4	100%
miR-966	4	100%
miR-970	4	98.2%
miR-973	4	100%
miR-974	4	100%
miR-978	4	100%
miR-982,303	4	100%
miR-983-1	4	76.5%
miR-984, 983-1, 983-2	4	100%
miR-986	4	100%
miR-987	4	100%
miR-988	4	100%
miR-992	4	100%
miR-993	4	100%
miR-994	4	100%
miR-9a	4	100%
wt/ldn	4	96.5%
W^{1118}/LDN	4	100%

The genetic screen was carried out by scoring the both sides of the halteres for the newly enclosed adult flies. An average of 50 flies, including half male and half female, are used for the genetic screen for each miRNAs. miRNA class denotes the classification of miRNAs according the morphological change of the haltere after the overexpression of the miRNAs in the haltere imaginal discs by the *Ubx*-$Gal4^{LDN}$.